



TESIS SS14-2501

PENARIKAN SAMPEL DENGAN BALANCED
SAMPLING DAN PENDEKATAN BAYESIAN UNTUK
SURVEI KETENAGAKERJAAN
STUDI KASUS: PROVINSI KALIMANTAN TENGAH

MOCH. CHOERIL ANWAR
NRP 1313 201 716

DOSEN PEMBIMBING
Prof. Drs. Nur Iriawan, M.Ikom., Ph.D.
Dr. Drs. Brodjol Sutijo Suprih Ulama, M.Si.

PROGRAM MAGISTER
JURUSAN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2015



THESIS SS14-2501

SAMPLING DESIGN WITH BALANCED SAMPLING
AND BAYESIAN APPROACH FOR
LABOUR FORCE SURVEY
CASE STUDY: CENTRAL KALIMANTAN PROVINCE

MOCH. CHOERIL ANWAR
NRP 1313 201 716

SUPERVISOR
Prof. Drs. Nur Iriawan, M.Ikom., Ph.D.
Dr. Drs. Brodjol Sutijo Suprih Ulama, M.Si.

MAGISTER PROGRAM
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS AND NATURAL SCIENCE
SEPULUH NOPEMBER INSTITUTE OF TECHNOLOGY
SURABAYA
2015

**PENARIKAN SAMPEL DENGAN *BALANCED SAMPLING*
DAN PENDEKATAN BAYESIAN UNTUK
SURVEI KETENAGAKERJAAN
STUDI KASUS: PROVINSI KALIMANTAN TENGAH**

**Tesis ini disusun untuk memenuhi salah satu syarat memperoleh gelar
Magister Sains (M.Si)
di
Institut Teknologi Sepuluh Nopember**

oleh:

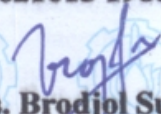
**MOCH. CHOERIL ANWAR
NRP. 1313 201 716**

**Tanggal Ujian : 5 Maret 2015
Periode Wisuda : September 2015**

Disetujui oleh:


**1. Prof. Drs. Nur Iriawan, M.Ikom, Ph.D
NIP. 19621015 198803 1 002**

(Pembimbing)


**2. Dr. Drs. Brodjol Sutijo Suprih Ulama, M.Si
NIP. 19660125 199002 1 001**

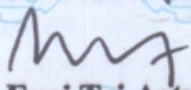
(Pembimbing)


**3. Dr. Irhamah, M.Si
NIP. 19780406 200112 2 002**

(Penguji)


**4. Dr. Kartika Fithriasari, M.Si
NIP. 19691212 199303 2 002**

(Penguji)


**5. Dr. Erni Tri Astuti, M.Math
NIP. 19671022 199003 2 002**

(Penguji)

Direktur Program Pascasarjana,


**Prof. Dr. Ir. Adi Soeprijanto, M.T
NIP. 19640405 199002 1 001**

PENARIKAN SAMPEL DENGAN *BALANCED SAMPLING* DAN PENDEKATAN BAYESIAN UNTUK SURVEI KETENAGAKERJAAN STUDI KASUS: PROVINSI KALIMANTAN TENGAH

Nama Mahasiswa : Moch. Choeril Anwar
NRP : 1313 201 716
Pembimbing : Prof. Drs. Nur Iriawan, M.Ikom., Ph.D.
Co-Pembimbing : Dr. Drs. Brodjol Sutijo Suprih Ulama, M.Si.

ABSTRAK

Data ketenagakerjaan dapat menggambarkan kondisi perekonomian dan sosial di suatu wilayah dalam kurun waktu tertentu. Badan Pusat Statistik (BPS) sebagai instansi yang bertanggung jawab dalam pengumpulan data ketenagakerjaan telah menggunakan desain sampling yang kompleks dalam Survei Angkatan Kerja Nasional (Sakernas). Namun, masih ada beberapa kritik mengenai data ketenagakerjaan. Selain itu, estimasi Tingkat Pengangguran Terbuka (TPT) di beberapa kabupaten dinilai kurang *robust* karena mengalami penurunan yang cukup tajam pada tahun-tahun terakhir. Hal ini kemungkinan diakibatkan oleh sampel blok sensus yang tidak merata di suatu kabupaten, yaitu antara daerah yang memiliki perekonomian yang baik dengan yang perekonomiannya tidak baik. Oleh karena itu, sangat penting dilakukan penelitian yang dapat memberikan kontribusi dalam sampling untuk survei ketenagakerjaan. Penelitian ini bertujuan memperoleh cara optimal penarikan sampel untuk survei ketenagakerjaan dengan pendekatan *balanced sampling* dan Bayesian, yaitu dengan melakukan estimasi parameter ketenagakerjaan dari sampel yang telah didapatkan dari kedua desain dan sampel dari desain yang digunakan BPS, dan membandingkan hasil estimasi dalam rangka mengetahui desain sampling yang paling baik. Metode sampling yang digunakan adalah *cube method* pada *balanced sampling* dan *reference prior* pada sampling dengan pendekatan Bayesian. Kerangka sampel yang digunakan berasal dari *raw* data Sensus Penduduk 2010 (SP2010). Penelitian ini telah berhasil membuat cara/ algoritma penarikan sampel dengan pendekatan Bayesian yaitu dengan *reference prior sampling*. Dari sampel BPS, dengan metode estimasi *design based*, diperoleh nilai absolut bias sebesar 0,0168, varians sebesar $1,96 \times 10^{-5}$, dan MSE sebesar $3,02 \times 10^{-4}$. Dengan metode Bayesian, diperoleh nilai absolut bias sebesar 25×10^{-5} , varians sebesar $6,67 \times 10^{-8}$, dan MSE sebesar $1,29 \times 10^{-7}$. Pada *balanced sampling* dengan *cube method*, dengan estimasi Horvitz-Thompson diperoleh nilai absolut bias sebesar 8×10^{-5} , varians sebesar $3,04 \times 10^{-6}$, dan MSE sebesar $3,05 \times 10^{-6}$. Sementara itu dengan metode Bayesian, diperoleh nilai absolut bias sebesar 9×10^{-5} , varians sebesar $6,83 \times 10^{-8}$, dan MSE sebesar $7,66 \times 10^{-8}$. Pada penarikan sampling dengan *reference prior sampling* ini diperoleh absolut bias sebesar 0,02, varians sebesar $6,96 \times 10^{-8}$, dan MSE sebesar $3,20 \times 10^{-6}$. Dengan hasil tersebut, diketahui bahwa desain *balanced sampling* merupakan desain yang lebih baik bila dibandingkan *reference prior sampling* maupun desain yang telah digunakan BPS didasarkan pada nilai MSE yang dihasilkan.

Kata Kunci : *balanced sampling*, *cube method*, Bayesian, *reference prior sampling*, akurasi, data ketenagakerjaan

SAMPLING DESIGN WITH BALANCED SAMPLING AND BAYESIAN APPROACH FOR LABOUR FORCE SURVEY CASE STUDY: CENTRAL KALIMANTAN PROVINCE

By : Moch. Choeril Anwar
Student Identity Number : 1313 201 716
Supervisor : Prof. Drs. Nur Iriawan, M.Ikom., Ph.D.
Co-Supervisor : Dr. Drs. Brodjol Sutijo Suprih Ulama, M.Si.

ABSTRACT

The employment data can describe the economic and social conditions, even the level of social welfare in an area and a certain time. BPS-Statistics Indonesia, the agency which responsible for the collection of employment data, has been using a complex sampling design in the National Labour Force Survey (Sakernas). However, there are some criticism regarding the employment data. In addition, the estimates of unemployment rate (TPT) in some districts considered less robust due to a sharp decline during last year. This is likely caused by census block samples uneven in some districts, between regions that have a good economy with the other. Therefore, it is very important to do research that can contribute for sampling design of employment survey. The objective of this research were to obtain the optimal sampling for employment survey with a balanced approach and Bayesian sampling, by estimating the parameters of employment from both sampling design and BPS's sampling design, and compare the results in order to determine which the best sampling design. The sampling method used was a cube method on balanced sampling method and the reference prior on sampling with Bayesian approach. The sampling frame were taken from the raw data of Population Census 2010 (SP2010). From this study, obtained sampling with Bayesian approach, that is reference prior sampling. From BPS's samples, with design-based estimation, the value of absolute bias, variance, and MSE was 0.0168 , $1,96 \times 10^{-5}$, and $3,02 \times 10^{-4}$. With Bayesian method, the value of absolute of bias, variance, and MSE was 25×10^{-5} , $6,67 \times 10^{-8}$, and $1,29 \times 10^{-7}$. From balanced sampling with the cube method and with the Horvitz-Thompson estimator, the value of absolute bias, variance, and MSE was 8×10^{-5} , $3,04 \times 10^{-6}$, and $3,05 \times 10^{-6}$. With Bayesian method, the value of absolute of bias, variance, and MSE was 15×10^{-5} , $6,85 \times 10^{-8}$, and $9,13 \times 10^{-8}$. On the reference priors sampling, the value of absolute bias, variance, and MSE was $0,02$, $6,96 \times 10^{-8}$, and $3,20 \times 10^{-6}$. With these results, it is known that balanced sampling design is better design than both the reference prior sampling and BPS design, based on the MSE value.

Keywords : balanced sampling, cube method, bayesian, reference prior sampling, accuracy, employment data

KATA PENGANTAR

Alhamdulillah Rabbil Aalamiin, puji syukur atas kehadiran Allah SWT, Dzat Yang Maha Esa, atas segala karunia nikmat, taufik dan hidayah-Nya kepada penulis sehingga tesis yang berjudul **“Penarikan Sampel dengan *Balanced Sampling* dan Pendekatan Bayesian untuk Survei Ketenagakerjaan, Studi Kasus: Provinsi Kalimantan Tengah”** dapat terselesaikan. Sholawat dan salam ditujukan kepada Nabi Muhammad SAW yang telah menuntun umat manusia dari kegelapan.

Pada kesempatan ini penulis ingin menyampaikan penghargaan dan ucapan terima kasih yang mendalam kepada:

1. Kepala Badan Pusat Statistik (BPS) RI, Kepala Pusdiklat BPS, Kepala BPS Provinsi Kalimantan Tengah dan Kepala BPS Kabupaten Sukamara yang telah memberi kesempatan serta beasiswa kepada penulis untuk melanjutkan studi program S2 di ITS.
2. Bapak Prof. Drs. H. Nur Iriawan, M.IKom, P.hD dan Bapak Dr. Drs. Brodjol Sutijo Suprih Ulama, M.Si atas segala bimbingan dan arahan dengan penuh kesabaran kepada penulis dalam penyusunan tesis ini.
3. Ibu Dr. Irhamah, M.Si, Ibu Dr. Kartika Fithriasari, M.Si, dan Ibu Dr. Erni Tri Astuti, M.Math selaku dosen penguji yang telah memberikan saran dan koreksi atas penulisan tesis ini.
4. Bapak Dr. Suhartono, M.Sc selaku Kepala Program Studi S2/S3 Jurusan Statistika FMIPA ITS beserta jajarannya yang telah memberikan fasilitas selama proses studi.
5. Bapak Ibu dosen Statistika ITS yang telah mencurahkan ilmu dan pengalamannya selama proses studi.
6. Ibunda Ulwiyah dan Almarhum Abah Slamet Sodikin serta seluruh keluarga besarku atas segala do'a dan dukungannya sehingga penulis berhasil menyelesaikan studi dengan baik.
7. Istriku tercinta, Dyan Ayu Dwi Agustin, terima kasih atas segala dukungan, do'a dan cintamu. Anak-anakku, permata hatiku, Muhammad

Alif Jundullah (3 th 8 bln) dan Raisa Saliha Malahati (11 bln). Do'a dan harapan terbaik selalu untuk kalian.

8. Mas Syahrul Khoir, terima kasih atas kebaikan hati dan bantuan ilmunya.
9. Teman-teman angkatan 7, Mas Ade, Bang Aal, Mas Cahyo, Bang Heri, Untung, Mas Nora, Gama, Hadi, Mbak Metti, Mbak Lilis, Mbak Ratna, Mbak Tika, Mbak Reni, Mbak May, Mbak Arifah, Rini, Mbak Maya, Mbak Eta, dan Mbak Devi. Bersyukur dapat bertemu dengan temen-temen semua. Semoga kita dapat berjumpa lagi di lain kesempatan.
10. Rekan-rekan di BPS Kalimantan Tengah, Mbak Ervin, Yansyah, dkk. atas bantuan data dan moral yang telah diberikan.
11. Semua pihak yang telah membantu penyelesaian tesis ini.

Akhirnya, do'a dan harapan selalu dipanjatkan kepada Allah SWT agar ilmu yang telah diperoleh menjadi barokah dan bermanfaat bagi sesama serta dapat menjadi sarana meraih ridho-Nya. Aamiin Ya Robbal 'Alamin.

Surabaya, Maret 2015

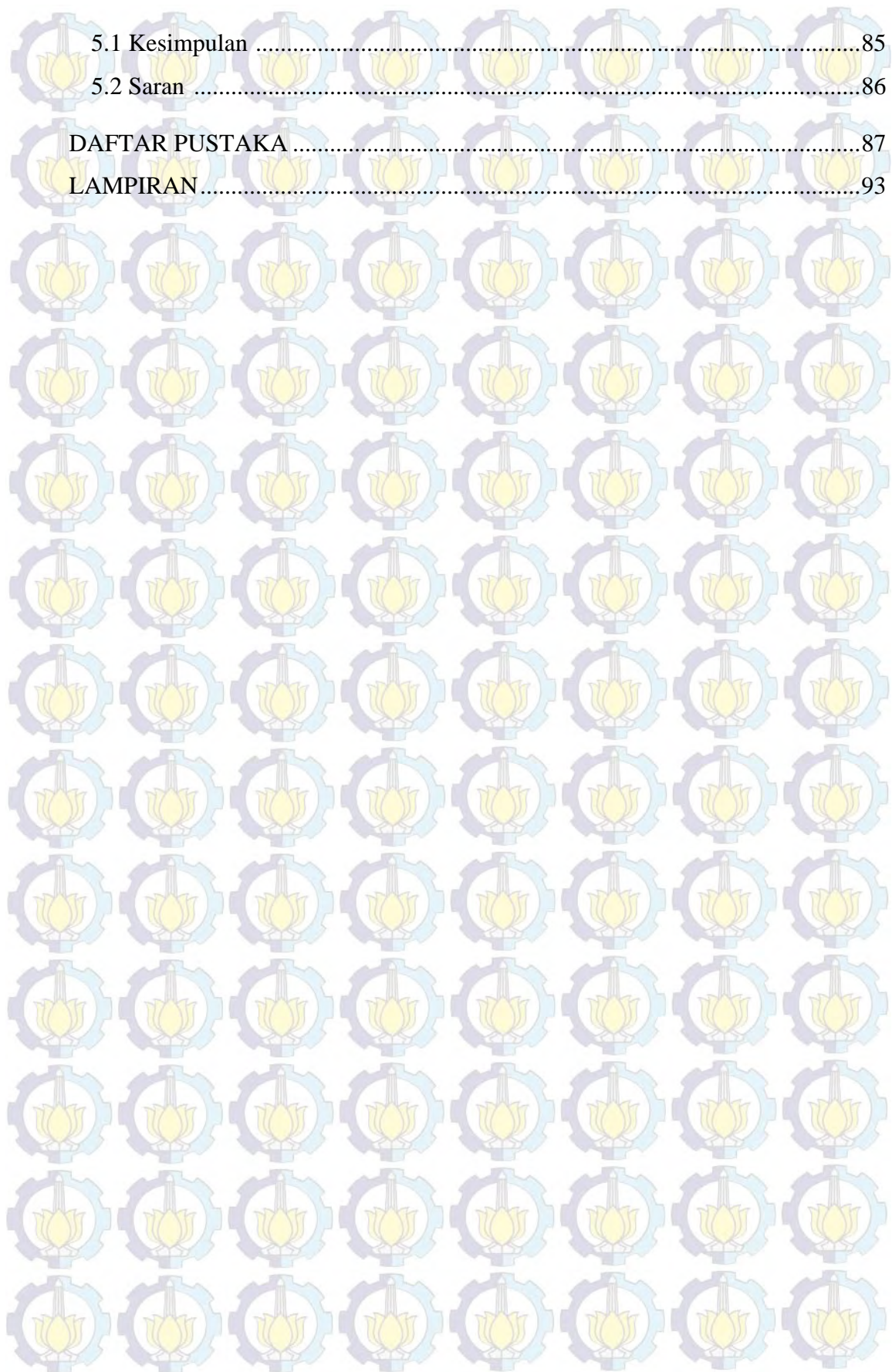
Penulis

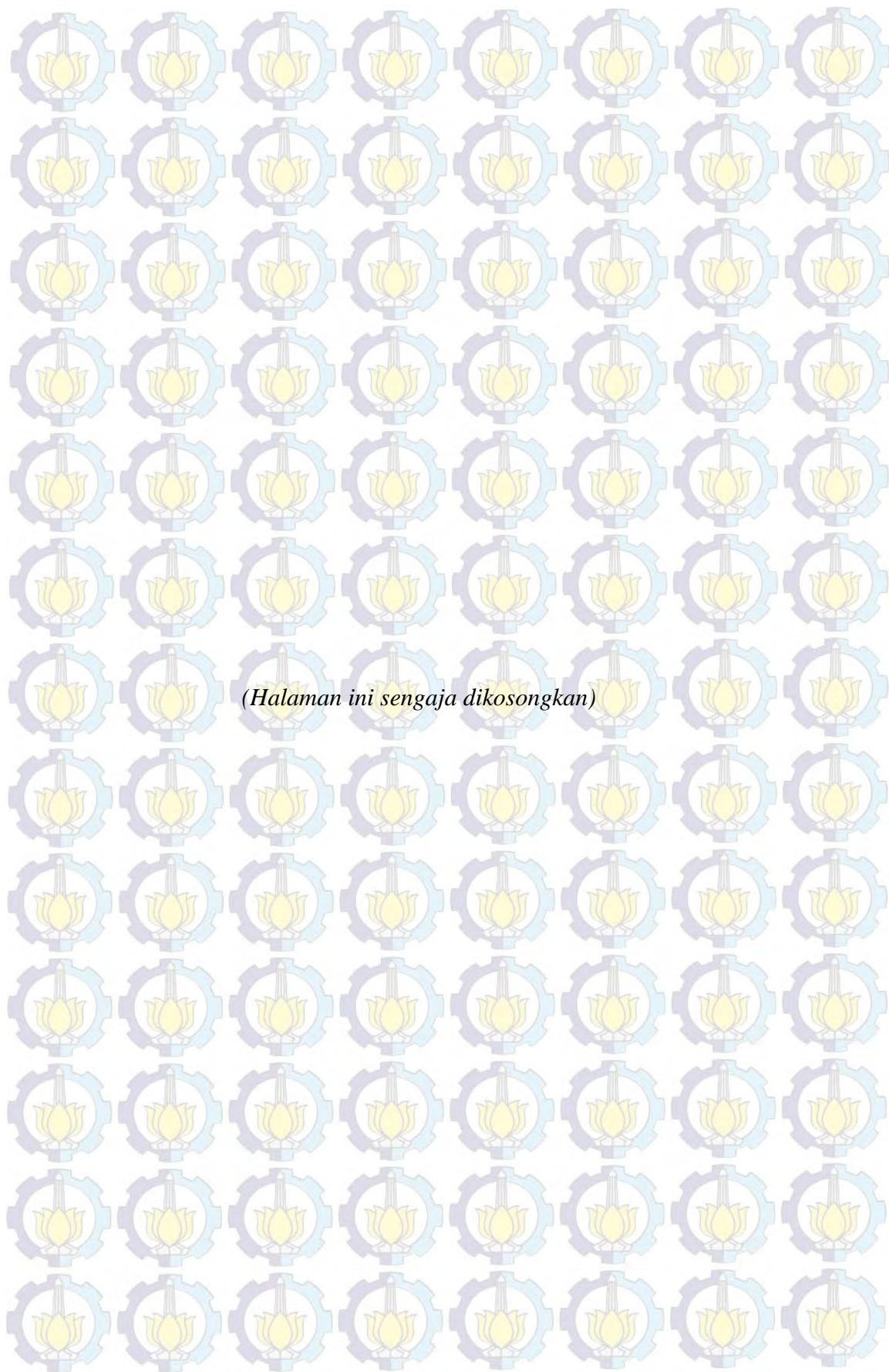
DAFTAR ISI

JUDUL PENELITIAN	i
LEMBAR PENGESAHAN	v
ABSTRAK	vii
ABSTRACT	viii
KATA PENGANTAR	ix
DAFTAR ISI	xi
DAFTAR TABEL	xv
DAFTAR GAMBAR	xvii
DAFTAR LAMPIRAN	xix
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	7
1.3 Tujuan Penelitian	8
1.4 Manfaat Penelitian	8
1.5 Batasan Permasalahan	8
BAB 2 TINJAUAN PUSTAKA	11
2.1 Kerepresentatifan (<i>Representativeness</i>)	11
2.2 Konsep Dasar Sensus dan Survei	14
2.3 Jenis Pengambilan Sampel	15
2.3.1 Sampling Tidak Berpeluang (<i>Non-probability Sampling</i>)	15
2.3.2 Sampling Berpeluang (<i>Probability Sampling</i>)	16
2.4 Desain Sampling Survei Angkatan Kerja Nasional (Sakernas) 2011	18
2.5 Estimasi Rasio dan Estimasi Varians Rasio Dua Nilai Karakteristik	21
2.6 Karakteristik Estimator dari Survei dan Desain Sampling yang Baik	22
2.7 <i>Balanced Sampling</i>	23
2.7.1 Definisi <i>Balanced Sampling</i>	25
2.7.2 Tujuan <i>Balanced Sampling</i>	27
2.7.3 <i>Cube Method</i>	27
2.7.4 Tahapan <i>Balanced Sampling</i> dengan <i>Cube Method</i>	29

2.8 Metode Bayesian	35
2.8.1 Distribusi Prior	36
2.8.2 <i>Hierarchical Bayes</i> dalam <i>Reference Prior Sampling</i>	37
2.8.3 <i>Markov Chain Monte Carlo</i>	39
2.8.4 <i>Gibbs Sampler</i>	41
2.8.5 Peran Bayesian dalam Survey Sampling	44
2.9 Ketenagakerjaan	48
2.9.1 Bekerja	48
2.8.2 Konsep Pengangguran dalam Sensus dan Survei BPS	49
2.8.3 Distribusi dari Data Ketenagakerjaan	50
BAB 3 METODOLOGI PENELITIAN	53
3.1 Sumber Data	53
3.2 Variabel Penelitian	53
3.3 Definisi Operasional	54
3.4 Struktur Data	55
3.5 Metode Penelitian	56
BAB 4 HASIL DAN PEMBAHASAN	63
4.1 Gambaran Umum Provinsi Kalimantan Tengah	63
4.2 Gambaran Kondisi Ketenagakerjaan Provinsi Kalimantan Tengah	
Berdasarkan Hasil Sensus Penduduk 2010	64
4.3 Karakteristik Data Ketenagakerjaan dari Sensus Penduduk 2010	68
4.4 Hasil Penarikan Sampel	69
4.4.1 Karakteristik Sampel Sakernas 2011 dari Desain Sampling BPS ...	69
4.4.2 Penarikan Sampel <i>Balanced Sampling</i> dengan <i>Cube Method</i>	70
4.4.3 Penarikan Sampel Pendekatan Bayesian dengan <i>Reference Prior</i> ..	71
4.4.3 Perbandingan Karakteristik Sampel yang Dihasilkan	73
4.5 Perbandingan Hasil Estimasi dari Ketiga Desain Sampling	75
4.5.1 Perbandingan Hasil Estimasi Proporsi Pengangguran	76
4.5.2 Perbandingan Hasil Estimasi Proporsi Pekerja menurut Lapangan	
Usaha	79
BAB 5 KESIMPULAN DAN SARAN	85

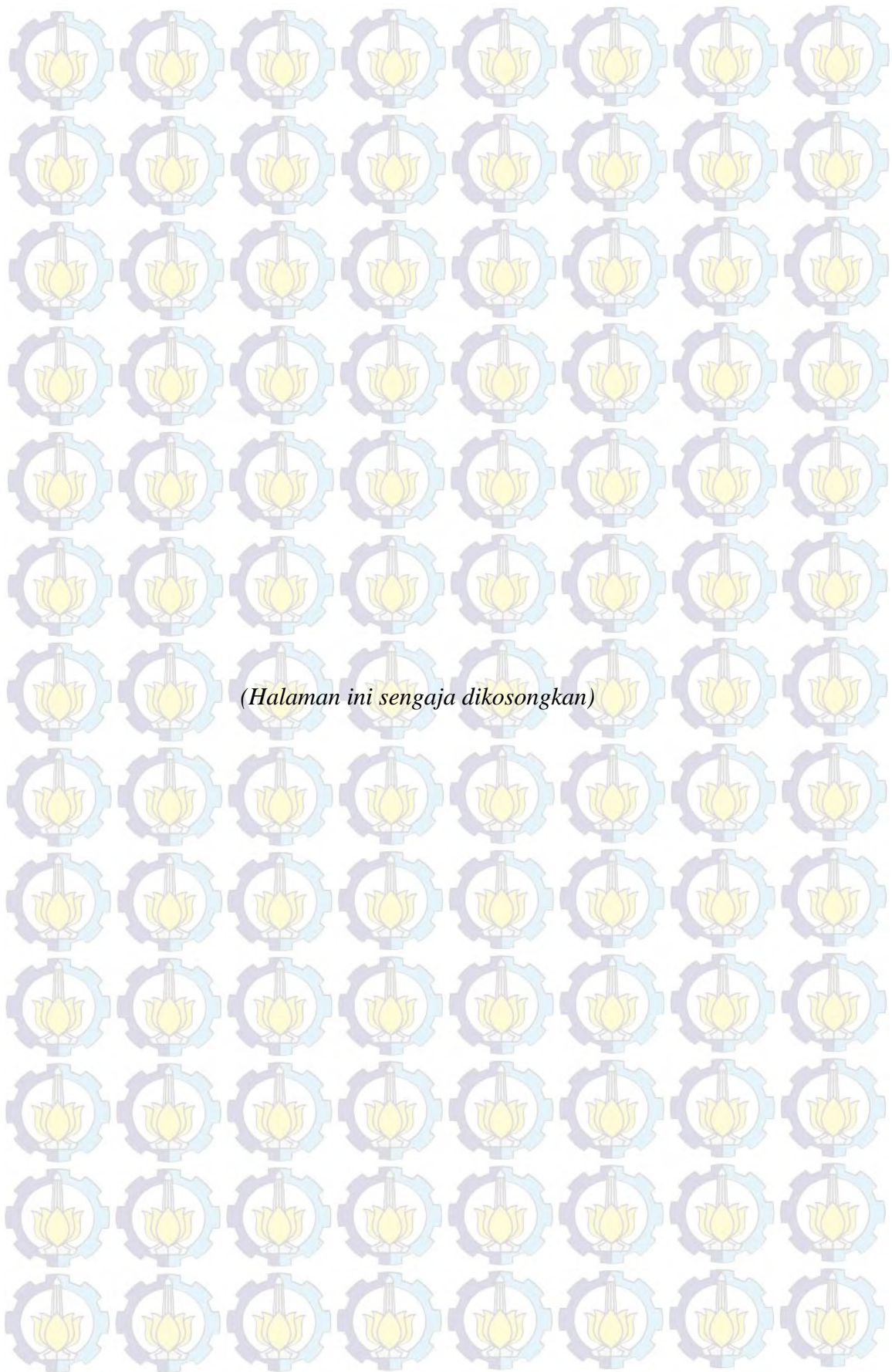
5.1 Kesimpulan	85
5.2 Saran	86
DAFTAR PUSTAKA	87
LAMPIRAN.....	93





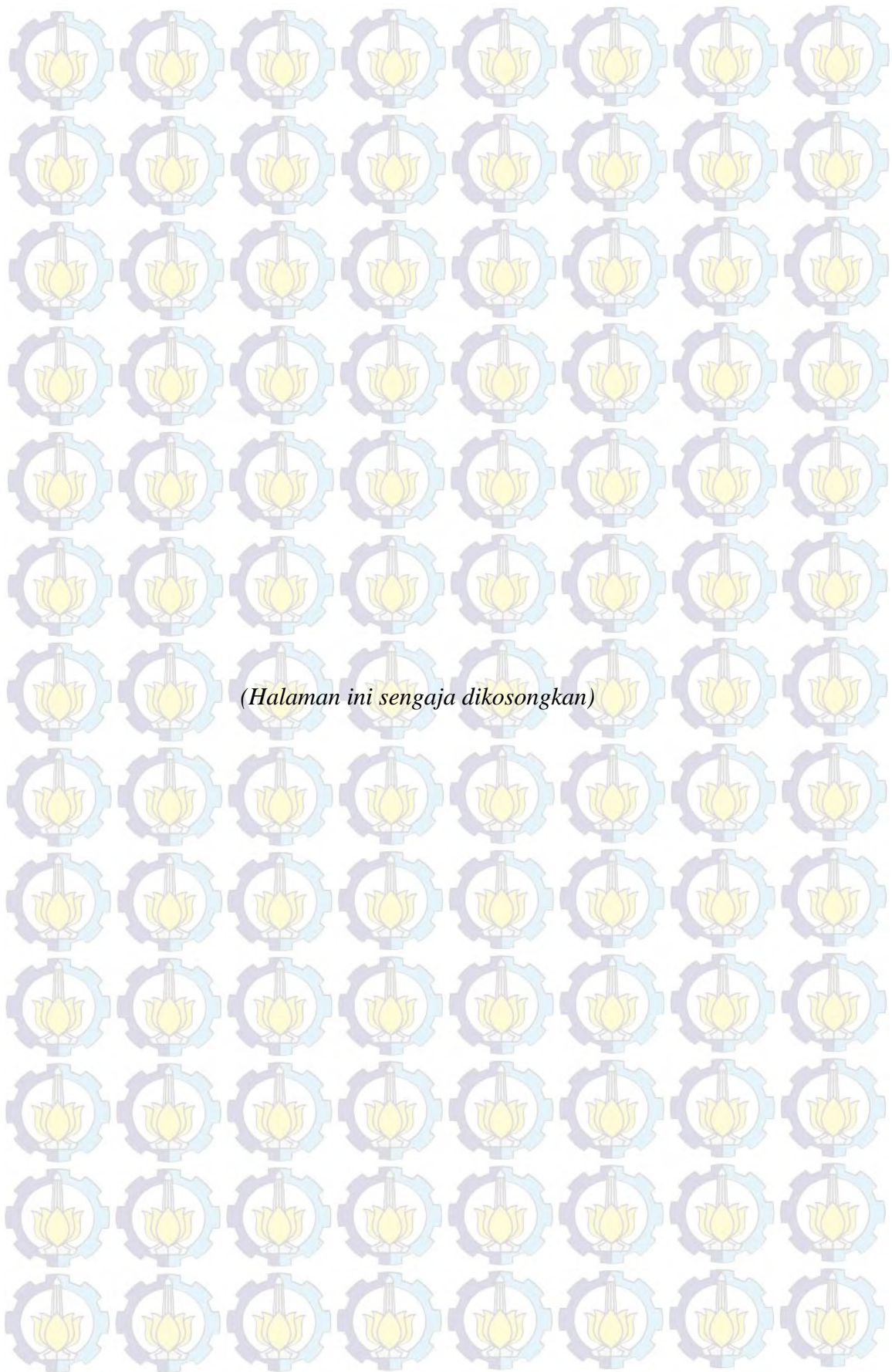
DAFTAR TABEL

Tabel 1.1	Tingkat Pengangguran Terbuka (TPT) Kalimantan Tengah periode 2008 s.d 2013.....	3
Tabel 2.1	Perbandingan Survei Sampel dan Sensus.....	14
Tabel 2.2	Desain Penarikan Sampel Sakernas 2011.....	19
Tabel 3.1	Struktur Data Penarikan Sampel dengan Pendekatan Bayesian (<i>Reference Prior</i>).....	55
Tabel 3.2	Struktur Data Penarikan Sampel dengan <i>Balanced Sampling</i> dengan <i>Cube Method</i>	56
Tabel 4.1	Hasil Uji Anderson Darling terhadap Data Pengangguran pada Populasi	68
Tabel 4.2	Karakteristik Data Blok Sensus di Kerangka Sampel	69
Tabel 4.3	Karakteristik Sampel dari Desain Sampling BPS.....	69
Tabel 4.4	Karakteristik Sampel dari <i>Balanced Sampling</i>	70
Tabel 4.5	Karakteristik Sampel dari <i>Reference Prior</i>	73
Tabel 4.6	Persebaran Sampel dan Karakteristik Sampel di 14 Kabupaten/ Kota	74
Tabel 4.7	Hasil Estimasi Jumlah Nilai Karakteristik dari Sampel Desain Sampling BPS dengan Metode <i>Design Based</i>	75
Tabel 4.8	Hasil Estimasi Jumlah Nilai Karakteristik dari Sampel <i>Balanced Sampling</i> dengan Metode Horvitz Thompson.....	76
Tabel 4.9	Perbandingan Estimasi Proporsi Pengangguran dengan Metode Estimasi Berbeda	77
Tabel 4.10	Perbandingan Estimasi Proporsi Pengangguran dengan Metode Estimasi Bayesian.....	78
Tabel 4.11	Perbandingan Estimasi Proporsi Pekerja menurut Lapangan Usaha dengan Metode Estimasi Berbeda	80
Tabel 4.12	Perbandingan Estimasi Proporsi Pekerja menurut Lapangan Usaha dengan Metode Estimasi Bayesian	81



DAFTAR GAMBAR

Gambar 2.1 (a) Representasi Geometrik dari Kemungkinan Sampel dari Populasi dengan $N=3$; (b) <i>Balanced Constrain</i> ukuran tetap dari jumlah sampel=2: sampel <i>balanced</i> ada; (c) <i>Balanced constrain</i> sedemikian sehingga sampel <i>balanced</i> tidak ada; (d) <i>Balanced Constrain</i> sedemikian sehingga sampel <i>balanced</i> ada yang dapat diwujudkan dalam sudut kubus dan ada yang tidak	29
Gambar 2.2 Skema Hirarki dalam <i>Reference Prior Sampling</i>	38
Gambar 2.3 Skema Penarikan Sampel dengan Strukur Hirarki dalam <i>Reference Prior Sampling</i>	39
Gambar 3.1 Diagram Ketenagakerjaan Sensus Penduduk 2010.....	57
Gambar 3.2 Diagram Alir Penelitian	61
Gambar 3.3 Diagram Alir <i>Balanced Sampling</i> dengan <i>Cube Method</i>	61
Gambar 3.4 Diagram Alir <i>Reference Prior Sampling</i>	61
Gambar 4.1 Piramida Penduduk Kalimantan Tengah Tahun 2010	64
Gambar 4.2 Diagram Pembagian Penduduk Usia 15 Tahun ke Atas	65
Gambar 4.3 Persentase Pengangguran terhadap Provinsi dan TPT Kabupaten/ Kota di Kalimantan Tengah.....	66
Gambar 4.4 Persentase Penduduk Kalimantan Tengah yang Bekerja menurut Kabupaten/ Kota dan Lapangan Usaha	67
Gambar 4.5 Perbandingan Nilai Estimasi Proporsi Pengangguran dengan Pengurangan dan Penambahan Standard Deviasi.....	79
Gambar 4.6 Perbandingan Nilai Estimasi Proporsi Pekerja menurut Lapangan Usaha dengan Penambahan dan Pengurangan Standard Deviasi: a. Pertanian, b. Industri, c. Perdagangan, rumah makan dan akomodasi, d. Jasa Kemasyarakatan dll	83



DAFTAR LAMPIRAN

Lampiran 1. Pertanyaan Ketenagakerjaan pada Kuesioner Sensus Penduduk 2010	93
Lampiran 2. Blok Sensus Sampel Sakernas Triwulanan 2011 Kalimantan Tengah dengan Data Variabel Penelitian dari SP2010.....	94
Lampiran 3. Syntax Program R Penarikan Sampel <i>Balanced Sampling</i> dengan <i>Cube Method</i>	98
Lampiran 4. Hasil Penarikan Sampel dengan Pendekatan <i>Balanced Sampling</i> dengan <i>Cube Method</i>	99
Lampiran 5. Blok Sensus dan Data Variabel Penelitian dari Sampel <i>Balanced Sampling</i> dengan <i>Cube Method</i>	102
Lampiran 6. Syntax Program R untuk Penarikan Sampel dengan Pendekatan <i>Reference Prior Sampling</i>	106
Lampiran 7. Hasil Penarikan Sampel dengan Pendekatan <i>Reference Prior</i>	108
Lampiran 8. Blok Sensus dan Data Variabel Penelitian dari Sampel <i>Reference Prior Sampling</i>	111
Lampiran 9. Syntax Program WinBUGS dan Hasil dari Estimasi Parameter Proporsi Pengangguran dari Sampel BPS	115
Lampiran 10. Syntax Program WinBUGS dan Hasil dari Estimasi Parameter Proporsi Pengangguran dari Sampel Hasil <i>Cube Method</i>	116
Lampiran 11. Syntax Program WinBUGS dan Hasil dari Estimasi Parameter Proporsi Pengangguran dari Sampel Hasil <i>Reference Prior</i>	117
Lampiran 12. Syntax Program WinBUGS Estimasi Parameter Proporsi Pekerja menurut Lapangan Usaha dari Sampel Sakernas BPS	118
Lampiran 13. Hasil Estimasi Parameter Proporsi Pekerja menurut Lapangan Usaha dari Sampel Sakernas BPS	119
Lampiran 14. Syntax Program WinBUGS Estimasi Parameter Proporsi Pekerja menurut Lapangan Usaha dari Sampel Hasil <i>Cube Method</i>	121
Lampiran 15. Hasil Estimasi Parameter Proporsi Pekerja menurut Lapangan Usaha dari Sampel Hasil <i>Cube Method</i>	122

Lampiran 16. Syntax Program WinBUGS Estimasi Parameter Proporsi Pekerja menurut Lapangan Usaha dari Sampel Hasil <i>Reference</i> <i>Prior</i>	124
Lampiran 17. Hasil Estimasi Parameter Proporsi Pekerja menurut Lapangan Usaha dari Sampel Hasil <i>Reference Prior</i>	125

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Salah satu indikator penting dalam pembangunan adalah data ketenagakerjaan. Data ini dapat menggambarkan kondisi perekonomian, sosial, bahkan tingkat kesejahteraan penduduk di suatu wilayah dan dalam suatu/ kurun waktu tertentu (BPS, 2011). Dari sisi ekonomi, data ketenagakerjaan ini dapat menggambarkan bagaimana kondisi perekonomian suatu wilayah. Salah satu data ketenagakerjaan adalah jumlah pengangguran atau yang sering digunakan yaitu Tingkat Pengangguran Terbuka (TPT). TPT yang meningkat dapat menggambarkan bahwa perekonomian sedang melambat. Dari sisi sosial, pengangguran yang meningkat akan berpotensi meningkatkan kerawanan sosial. Selain itu, jumlah penganggur intelektual yang meningkat juga menunjukkan bahwa pendidikan yang diselenggarakan tidak sesuai dengan kebutuhan dunia kerja. Data ketenagakerjaan juga dapat menggambarkan seberapa besar *entrepreneur* yang berkembang di masyarakat dan berapa banyak yang menjadi pegawai dan karyawan.

Instansi pemerintah yang bertanggung jawab dalam pengumpulan data ketenagakerjaan adalah Badan Pusat Statistik (BPS). BPS (2011) menjelaskan bahwa pengumpulan data ketenagakerjaan telah dilakukan melalui sensus dan survei, antara lain: Sensus Penduduk (SP), Survei Penduduk Antar Sensus (SUPAS), Survei Sosial Ekonomi Nasional (Susenas) dan Survei Angkatan Kerja Nasional (Sakernas). Dari beberapa kegiatan tersebut, Sakernas merupakan kegiatan yang dirancang khusus untuk mengumpulkan data yang dapat menggambarkan keadaan umum ketenagakerjaan antar periode pencacahan. Sakernas pertama kali dilaksanakan di Indonesia tahun 1976. Dalam perkembangannya Sakernas mengalami berbagai perubahan baik dalam periode pencacahan maupun cakupan sampel wilayah dan rumah tangga. Tahun 1986 sampai dengan 1993 Sakernas dilaksanakan secara triwulanan, tahun 1994 sampai dengan 2001 secara tahunan setiap bulan Agustus, sedangkan tahun 2002 sampai

dengan 2004 selain secara tahunan juga dilaksanakan secara triwulanan. Mulai tahun 2005 sampai dengan tahun 2010 Sakernas dilakukan secara semesteran. Pada tahun 2011 sampai dengan 2015 pendataan Sakernas kembali dilakukan kembali secara triwulanan yaitu; bulan Februari (Triwulan I), Mei (Triwulan II), Agustus (Triwulan III), dan November (Triwulan IV) yang penyajian data dirancang sampai tingkat provinsi. Pelaksanaan Sakernas Triwulan III (bulan Agustus) selain sampel triwulanan juga terdapat sampel tambahan, untuk kepentingan angka tahunan sebagai estimasi penyajian data sampai tingkat kabupaten/ kota.

Ada beberapa kritik tentang data ketenagakerjaan yang merupakan output Sakernas. Sebagai contoh, pada September 2014 Ketua Umum Asosiasi Pengusaha Indonesia (APINDO) mengatakan bahwa dunia usaha mempertanyakan BPS yang menyatakan jumlah pengangguran terbuka turun 50.000 seiring perlambatan ekonomi. Kondisi ini dinilai tidak logis karena penurunan pertumbuhan mencerminkan aktivitas bisnis yang melambat sehingga sulit menciptakan lapangan kerja baru. Sejalan dengan Ketua Apindo, peneliti Indef menambahkan bahwa sektor yang menyerap tenaga kerja, yaitu sektor industri padat karya dan sektor pertanian, mengalami perlambatan. Jadi, penurunan pengangguran yang kurang sesuai dengan kondisi pertumbuhan ekonomi ini menjadi sebuah pertanyaan (Neraca, 2014).

Pada estimasi Tingkat Pengangguran Terbuka (TPT) untuk tingkat kabupaten/ kota, TPT beberapa kabupaten/ kota dinilai tidak *robust*. Hal ini terjadi bersamaan dengan pergantian paket sampel yang baru mulai tahun 2011. Beberapa kabupaten/ kota mengalami penurunan yang cukup tajam hingga sekitar 50 persen. Beberapa kabupaten/ kota tersebut adalah Kabupaten Kotawaringin Timur, Kabupaten Barito Selatan, Kabupaten Barito Utara, Kabupaten Sukamara dan Kota Palangka Raya. Penurunan yang cukup tinggi di beberapa kabupaten/ kota berpengaruh pada penurunan TPT Provinsi Kalimantan Tengah yang cukup tinggi pula, yaitu dari 4,14 persen menjadi 2,55 persen pada tahun 2011. Sebagai contoh lebih rinci adalah Kabupaten Sukamara, Kalimantan Tengah. Pada periode sebelum tahun 2011, TPT berkisar pada angka lima persen, tetapi mulai tahun 2011 TPT berkisar pada angka kurang dari satu persen.

Penurunan yang cukup tajam ini mengakibatkan TPT pada periode 2011 s.d 2013 persentasenya aneh. Hal ini juga menjadi pertanyaan bagi pemerintah daerah/ *stakeholder* di Kabupaten Sukamara. Selain Kabupaten Sukamara, terdapat empat kabupaten/ kota yang mengalami penurunan TPT yang cukup tajam, yaitu Kabupaten Kotawaringin Timur, Kabupaten Barito utara, dan Kota Palangka Raya. Pada Tabel 1.1 ditampilkan *series* data TPT kabupaten/ kota di Kalimantan Tengah 2008 s.d 2013.

Tabel 1.1 Tingkat Pengangguran Terbuka (TPT) Kalimantan Tengah periode 2008 s.d 2013

Kabupaten/Kota	2008	2009	2010	2011	2012	2013
1. Kotawaringin Barat	5,43	4,74	4,46	2,81	2,36	3,68
2. Kotawaringin Timur	4,48	4,42	4,16	2,08	4,61	2,45
3. Kapuas	4,40	5,35	4,17	3,52	3,08	1,51
4. Barito Selatan	6,30	4,00	5,51	2,52	1,73	2,17
5. Barito Utara	3,43	3,50	2,30	0,70	1,63	3,01
6. Sukamara	4,13	5,66	5,32	0,58	0,32	1,85
7. Lamandau	5,41	4,86	2,95	2,53	0,89	1,67
8. Seruyan	3,36	3,52	4,46	2,41	3,99	4,65
9. Katingan	3,42	4,42	3,34	2,61	3,39	5,82
10. Pulang Pisau	3,72	2,26	2,11	2,62	2,59	2,40
11. Gunung Mas	3,35	4,15	4,45	2,38	4,41	3,00
12. Barito Timur	3,46	3,22	1,35	2,07	0,95	1,68
13. Murung Raya	2,51	1,62	1,03	1,32	1,52	3,54
71. Palangka Raya	8,20	9,17	8,48	3,82	6,38	5,12
Kalimantan Tengah	4,59	4,62	4,14	2,55	3,17	3,09

Sumber: Hasil Sakernas

Rendahnya TPT Sukamara dan beberapa kabupaten/ kota lain pada tahun 2011 s.d 2013 dapat disebabkan oleh beberapa kemungkinan. Pertama, perekonomian yang baik di suatu wilayah akan menyebabkan penyerapan tenaga kerja dan pada akhirnya dapat menurunkan tingkat pengangguran terbuka. Kedua, blok sensus yang telah ditetapkan BPS sebagai sampel Sakernas pada periode 2011 s.d 2013 kurang merata antara daerah yang memiliki potensi ekonomi yang baik dengan yang tidak. Kondisi ini memungkinkan kecamatan yang memiliki

jumlah pengangguran yang cukup besar tidak terwakili di Sakernas periode 2011 s.d 2013.

Desain sampling yang digunakan mulai Sakernas 2011 adalah penarikan sampel tiga tahap berstrata (*Stratified Three Stage Sampling*). Populasi dibagi menjadi dua strata, yaitu perkotaan dan perdesaan. Penarikan sampel dilakukan independen untuk masing-masing strata. Pada tahap pertama, penarikan sampel wilayah cacah secara *probability proportional to size* (pps) dengan *size* jumlah rumah tangga SP2010. Tahap kedua menarik sampel blok sensus secara pps sistematis dan tahap ketiga menarik sampel rumah tangga secara sistematis (BPS, 2012). Jadi, desain sampling yang digunakan di dalam Sakernas merupakan desain yang kompleks. Desain yang kompleks terbukti telah memberikan desain yang murah dan efektif dalam mendapatkan informasi. Desain kompleks seperti ini juga digunakan oleh berbagai negara (Murphy, 2008). Desain sampling bertahap memiliki beberapa keunggulan, yaitu: tidak memerlukan kerangka sampel sampai unit terkecil, lebih efisien dalam biaya, tenaga, dan waktu serta dapat mengurangi *non-sampling error*. Namun demikian, metode sampling bertahap kurang akurat bila dibandingkan dengan *simple random sampling* (SRS). Desain sampling berstrata memiliki keunggulan sebagai berikut: masing-masing strata dapat diperlakukan sebagai populasi tersendiri, membawa manfaat pada administrasi, setiap strata dapat diterapkan rancangan sampling yang berbeda, dan dapat diperoleh estimasi dengan presisi yang lebih tinggi (Cochran, 1977). Opsomer (2011) menjelaskan bahwa stratifikasi kurang fleksibel karena Stratifikasi membagi populasi ke dalam sel/ bagian yang didefinisikan oleh variabel stratifikasi, yang mungkin menyebabkan proliferasi banyak sel kecil. Legg dan Yu (2010) menyampaikan bahwa pada sampling acak berstrata kemungkinan terjadi sampel yang tidak diinginkan.

Meskipun desain sampling yang digunakan BPS merupakan desain sampling yang kompleks dan memiliki berbagai keunggulan, tetapi hasil yang diperoleh di beberapa wilayah masih dipertanyakan. Hal ini mendorong penelitian untuk mempelajari desain sampling lain yang telah berkembang dengan keunggulannya masing-masing. Desain lain tersebut diharapkan dapat

memberikan alternatif baru yang mungkin dipertimbangkan BPS dalam desain sampling.

Tille (2006) menyampaikan bahwa sampling merupakan proses memilih unit dari populasi sehingga dari sampel dimungkinkan mengestimasi parameter populasi yang tidak diketahui. Istilah sampling mulai digunakan oleh Tschuprow pada 1923 dan Neyman pada 1934. Namun demikian, paper pertama tentang prosedur sampling adalah *unequal probability sampling*. Hansen dan Hurwitz (1943) dalam Tille (2006) mengusulkan desain multinomial *with replacement* dengan sebuah estimator tidak bias. Kemudian berkembang *unequal probability sampling without replacement* yang jauh lebih rumit, tetapi sebagian besar metode yang diusulkan dibatasi untuk ukuran sampel sama dengan dua. Brewer dan Hanif (1983) dalam Tille (2006) membangun 50 metode sampling, tetapi hanya 20 yang benar-benar bekerja dan banyak prosedur lain yang sangat lambat untuk diterapkan. Desain yang terkenal adalah desain sampling sistematis. Setelah tahun 1983, banyak metode sampling baru dengan *unequal probability* dipublikasikan.

Tille (2006) membahas beberapa metode yang telah berkembang. Salah satu yang berkembang adalah *balanced sampling*. Jika nilai rata-rata dari beberapa variabel diketahui untuk setiap unit dari seluruh populasi, maka dimungkinkan untuk memperoleh sampel yang rata-ratanya paling mendekati dengan rata-rata populasi, pada kasus ini dinamakan *balanced sample* (Yates, 1949). Royall dan Herson (1973) membahas pentingnya *balanced sampling* untuk melindungi kesimpulan dari *misspecified model*. Mereka mengusulkan estimator yang optimal dengan pendekatan model regresi. Dalam pendekatan berbasis model ini, optimalitas yang dikandung hanya berkenaan dengan model regresi tanpa memperhitungkan desain sampling, sehingga sampel harus seimbang (*balance*), tetapi tidak harus acak. Beberapa penelitian mengenai metode *balanced sampling* dalam Tille (2006) dilakukan antara lain oleh Yates (1946), Thionet (1953), Deville dkk. (1988), Ardilly (1991), Deville (1992), Hedayat dan Majumdar (1995), Deville dan Tille (1998), Valliant dkk. (2000), dan Deville dan Tille (2004).

Deville dan Tille (1998) mengusulkan delapan metode baru, salah satu di antaranya *splitting method*. Metode ini merupakan cara untuk menyajikan dengan

lebih sederhana beberapa metode *balanced sampling* yang telah ada. Metode ini kemudian dikembangkan lagi oleh Deville dan Tille (2004) menjadi *cube method*. *Cube method* memungkinkan pemilihan sampel seimbang acak pada beberapa variabel *balancing* dengan peluang terpilih sama atau tidak sama. Dalam arti bahwa estimator Horvitz-Thompson adalah sama atau hampir sama dengan populasi total variabel *balancing*. Deville dan Tille (2005) juga telah mengusulkan perkiraan varians untuk estimator Horvitz-Thompson pada *balanced sampling* dengan entropi besar.

Opsomer (2011) menjelaskan bahwa setidaknya ada dua tujuan *balanced sampling*, yaitu memperoleh representasi struktur populasi pada sampel dan meningkatkan efisiensi dari estimator survey. *Balanced sampling* dapat dilihat sebagai generalisasi dari stratifikasi, tetapi dapat menggunakan variabel apapun, baik variabel kategorik, variabel kontinu, ataupun kombinasi dari keduanya. Hal ini memberikan fleksibilitas dalam konstruksi sampel.

Pada analisis Bayesian, desain sampling bukan merupakan hal yang perlu diistimewakan. Hal ini disebabkan distribusi posterior dari parameter yang diestimasi diperlakukan sebagai variabel dan bersifat independen dengan bagaimana sampel dipilih (Basu, 1968). Namun, Zack (1969) menjelaskan bahwa set data yang berbeda kemungkinan akan membawa inferensi Bayesian pada tingkat presisi yang berbeda. Oleh karena itu, dia melakukan pemilihan sampel yang optimal dengan jumlah sampel yang telah ditentukan. Desain Bayesian yang dilakukannya bersifat tidak *random*, tanpa pengembalian, dan dilakukan secara sekuensial. Desain ini berusaha untuk meningkatkan presisi dengan meminimumkan *expected Bayes risk*. Meeden (2012) menyampaikan bahwa desain Bayesian ini belum pernah dipraktekkan. Permasalahan berada dalam ketepatan untuk menemukan distribusi prior yang bersesuaian dengan informasi prior dari populasi yang diamati.

Tujuan penggunaan sampling dengan pendekatan Bayesian dengan adalah untuk memperoleh representasi dari populasi. Hal ini dikarenakan sampling dengan pendekatan Bayesian mempertimbangkan peluang distribusi terpilih antar wilayah berdasarkan suatu variabel yang sesuai dengan survey yang akan

dilakukan. Dengan demikian, diharapkan pula efisiensi dari estimasi yang nanti diperoleh.

Penelitian ini akan berusaha melakukan sampling dengan desain *Balanced sampling* dan desain Bayesian. Kemudian kedua metode ini akan dibandingkan tingkat akurasi dengan desain sampling yang telah digunakan oleh BPS. Penelitian yang membandingkan beberapa metode sampling merupakan penelitian yang jarang dilakukan. Widaningsih (2006) melakukan perbandingan beberapa metode sampling dengan menduga populasi sapi potong di Kabupaten Karangasem. Metode sampling yang dibandingkan adalah acak sederhana, sistematis, acak berlapis, dan gerombol dua tahap. Hasil perbandingan menunjukkan teknik acak berlapis memberikan hasil estimasi dengan tingkat reliabilitas, validitas dan akurasi paling tinggi. Kusmayadi (2011) melakukan perbandingan metode sampling dengan menduga hasil Pemilukada Kabupaten Jembrana pada tahun 2010. Hasil perbandingan menunjukkan bahwa teknik penarikan acak berlapis (kecamatan dan desa-kota sebagai dasar) memberikan nilai dugaan dengan tingkat reliabilitas dan akurasi yang paling tinggi bila dibandingkan dengan teknik acak sederhana, teknik acak sistematis, dan teknik acak gerombol 2 tahap. Legg dan Yu (2010) melakukan perbandingan terhadap *rejection sampling* dan *cube sampling* dan keduanya menghasilkan estimator yang hampir sama (*equal*).

1.2 Rumusan Masalah

1. Bagaimana algoritma penarikan sampel untuk survei ketenagakerjaan dengan sampling pendekatan Bayesian.
2. Bagaimana penarikan sampel untuk survei ketenagakerjaan dengan *Balanced Sampling* dan sampling pendekatan Bayesian.
3. Bagaimana mengestimasi parameter distribusi variabel ketenagakerjaan dari sampel blok sensus BPS, *Balanced Sampling*, dan sampling pendekatan Bayesian dengan metode Bayesian.
4. Bagaimana perbandingan desain sampling yang telah digunakan di BPS dengan *Balanced Sampling* dan sampling pendekatan Bayesian dengan menggunakan varians, bias, dan MSE.

1.3 Tujuan Penelitian

1. Mendapatkan cara optimal penarikan sampel untuk survei ketenagakerjaan dengan sampling pendekatan *Bayesian*.
2. Mendapatkan sampel (blok sensus) untuk survei ketenagakerjaan dari *Balanced Sampling* dan sampling pendekatan Bayesian.
3. Mendapatkan estimasi parameter distribusi variabel ketenagakerjaan dari sampel blok sensus BPS, *Balanced Sampling*, dan sampling pendekatan Bayesian dengan metode Bayesian.
4. Mengetahui metode sampling yang lebih baik di antara desain sampling yang telah digunakan di BPS, *Balanced Sampling*, dan sampling pendekatan Bayesian dengan menggunakan varians, bias, dan MSE.

1.4 Manfaat Penelitian

1. Kontribusi dalam bidang keilmuan dalam menerapkan sampling pendekatan Bayesian dan *Balanced Sampling* untuk survey ketenagakerjaan.
2. Hasil penelitian diharapkan dapat menjadi bahan pertimbangan BPS dalam mendesain survey sampling.

1.5 Batasan Permasalahan

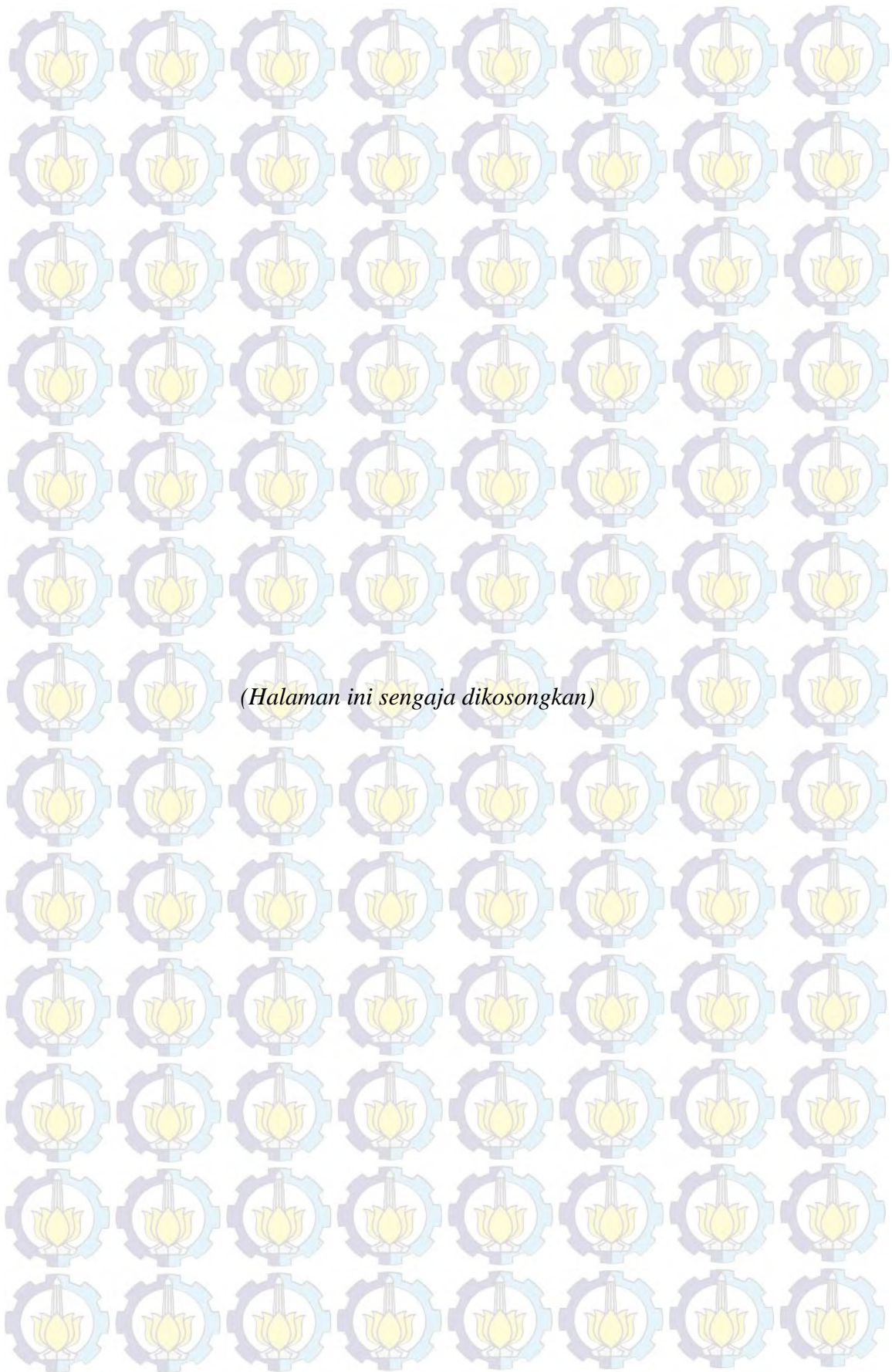
Dalam penelitian ini ruang lingkup permasalahan dibatasi pada penarikan sampel di Provinsi Kalimantan Tengah. Level Sakernas adalah kegiatan tingkat nasional yang pengambilan sampelnya dilakukan di BPS Pusat untuk seluruh wilayah Indonesia. Dalam penelitian ini, Provinsi Kalimantan Tengah dianggap sebagai *pilot project*.

Dalam rangka memberikan alternatif bagi BPS untuk mendesain sampling, penelitian ini hanya akan melakukan penarikan sampel *Balanced Sampling* dengan *cube method* dan penarikan sampel pendekatan Bayesian dengan *reference prior*.

Dengan beragamnya data ketenagakerjaan, penelitian ini dibatasi pada Tingkat Pengangguran Terbuka (TPT) dengan proksi proporsi pengangguran serta juga akan diteliti proporsi pekerja menurut lapangan usahanya. Untuk

mengestimasi parameter tersebut digunakan metode Bayesian dengan MCMC dan Gibbs Sampler pada sampel dari ketiga desain sampling, yaitu BPS, *Balanced Sampling*, dan sampling pendekatan Bayesian. Pada sampel dari BPS dan *Balanced Sampling* juga dilakukan estimasi total dari variabel untuk kemudian dihitung estimasi proporsi pengangguran dan proporsi pekerja menurut lapangan usaha. Kemudian dilakukan perbandingan di antara desain sampling yang telah digunakan di BPS, *Balanced Sampling*, dan sampling pendekatan Bayesian dengan menggunakan bias, varians, dan MSE.

Pada estimasi total dari suatu variabel dari sampel BPS, rumus estimasi dari desain *stratified three stage* disederhanakan menjadi *stratified one stage*. Hal ini dilakukan karena dua hal. Pertama, unit sampel pertama pada Sakernas yang blok sampelnya digunakan pada penelitian ini adalah rumah tangga, sedangkan pada penelitian ini unit sampel yang digunakan adalah blok sensus. Informasi tiap blok sensus sudah didapatkan secara langsung dari data SP2010 sehingga tidak memerlukan estimasi pada tingkat blok sensus dari rumah tangga terpilih. Kedua, tidak tersedianya data wilayah cacah yang bersesuaian dengan blok sensus pada *raw data* SP2010.



BAB 2

TINJAUAN PUSTAKA

2.1 Kerepresentatifan (*Representativeness*)

Langel dan Tille (2011) menerangkan bahwa ide dan konsep representatif dari sebuah populasi telah diperkenalkan oleh Kiaer pada tahun 1896, 1899, 1903, 1905 di pertemuan *International Statistical Institute* (ISI). Idennya mengenai sampel yang representatif meyakinkan orang dengan menyediakan gambaran validitas ilmiah. Representatif telah menjadi gagasan penting dalam pengambilan sampel sejak saat itu. Namun, banyaknya definisi representatif saat itu menimbulkan banyak perdebatan dan kesalahpahaman dalam sejarah sampling. Oleh karena itu, istilah ini sedikit digunakan dalam literatur survei sampling modern.

a. Kerepresentatifan dari Sampel

Langel dan Tille (2011) menyebutkan bahwa sifat kerepresentatifan objek statistik merupakan suatu hal yang jarang diklarifikasi. Hal ini dapat berupa unit statistik yang menjadi sampel (individu, daerah atau perusahaan yang representatif) atau strategi yang digunakan untuk prosedur pemilihan sampel. Penggunaan metode sampling yang memberikan sampel yang representatif dan menganalisis tingkat representatif dari sampel yang didapat merupakan dua konsep yang sangat berbeda. Dengan mempertimbangkan hal tersebut dan juga konsep keacakan dan kerepresentatifan, Hajek (1959, 1981) dalam Langel dan Tille (2011) menyebut istilah strategi yang representatif (tidak sampel yang representatif) ketika berbicara tentang metode pengambilan sampel seperti *balancing* atau kalibrasi, yaitu strategi yang terdiri atas sebuah desain sampling dan sebuah estimator. Namun, kemudian istilah kerepresentatifan ini lebih sering digunakan pada sampel.

b. Representatif sebagai miniatur dari populasi

Dalam banyak kasus, ide kerepresentatifan yang dipertimbangkan yaitu suatu sampel harus merupakan miniatur dari populasi. Hal ini merupakan

bentuk teoritis, yang menciptakan banyak divergensi dalam aplikasinya. Jika sampel yang dipilih akan menjadi replika populasi maka kompleksitas seluruh struktur populasi harus diketahui. Hal ini menjadi sesuatu yang tidak mungkin karena umumnya hanya aspek kecil yang diketahui. Selain itu, informasi lebih lanjut dalam populasi akan berbeda masing-masing unitnya. Pada akhirnya, sampel representatif yang sempurna adalah populasi itu sendiri. Hasil pengujian dari Neyman (1934) dalam stratifikasi menunjukkan bahwa mengharapkan sampel menjadi miniatur dari populasi yang ideal merupakan sebuah kekeliruan.

Dalam prakteknya, hanya mungkin untuk mendekati pandangan representatif dan banyak solusi yang berbeda telah diusulkan untuk mencapai itu. Metode ini tergantung pada pilihan *purposive* atau *random*, tetapi juga pada interpretasi yang berbeda dari kerepresentatifan.

c. Kekuatan Selektif

Kekuatan selektif dipandang sebagai jalan untuk kerepresentatifan. Hal ini dapat digambarkan oleh gagasan bahwa semakin sampel menyajikan karakter yang sama sebagaimana hal yang dikehendaki dari populasi, semakin dapat dinyatakan representatif. Representatif yang sempurna memang tidak didapatkan di dalam praktek, tetapi tetap bisa dipahami sebagai miniatur dari populasi dan didekati dengan menggunakan informasi yang tersedia sebagai kontrol.

Galvani (1951) dalam Langel dan Tille (2011) membedakan tiga prosedur sampling: pemilihan *random*, pemilihan *purposive* dan pemilihan bertingkat. Dalam makalahnya, pemilihan *random* adalah apa yang lebih sering dikenal sebagai *simple random sampling*. Ini tidak menggunakan informasi tambahan, dan karena itu tidak memerlukan pengetahuan sebelumnya dari populasi yang akan dijadikan sampel. Menurut Galvani, pernyataan ini tidak berarti bahwa setiap sampel yang didapat dari desain *simple random sampling* secara efektif mewakili semua karakteristik populasi.

Pemilihan *purposive*, sebaliknya, memerlukan informasi tambahan tentang seluruh populasi yang digunakan untuk menjamin kerepresentatifan

sampel dalam kaitannya dengan variabel-variabel tambahan. Namun, Galvani mencatat bahwa, tidak seperti dalam pemilihan acak, prosedur *purposive* tidak menjadi representatif untuk karakteristik yang tidak digunakan dalam proses seleksi. Galvani menganggap di satu sisi *simple random sampling* yang merupakan sebuah metode yang mewakili semua karakter dalam populasi, dan di sisi lain, pilihan *purposive* yang menghasilkan sampel yang representatif hanya untuk sejumlah karakteristik tertutup.

Stratified sampling dianggap oleh Galvani sebagai prosedur pemilihan *random* yang memanfaatkan pengetahuan tentang heterogenitas populasi mengenai beberapa karakteristik. Sifat ketidakberpihakan atau pengurangan totalitas juga dapat diberikan kepada *stratified sampling*, dengan mungkin akurasi yang lebih baik daripada *simple random sampling*. Dengan kata lain, Galvani berpikir bahwa stratifikasi bukan tipe kekuatan selektif yang dapat membahayakan kerepresentatifan. Dia menyatakan bahwa metode pemilihan *random* harus lebih dipilih karena memenuhi kondisi kerepresentatifan. Selain itu, teori probabilitas dapat diaplikasikan pada kasus *random*.

d. Representatif sebagai tujuan

Kruskal dan Mosteller (1979, 1980) dalam Langel dan Tille (2011) menekankan untuk tidak mempertanyakan keharusan kerepresentatifan menjadi tujuan survei sampling. Langel dan Tille (2011) menyatakan tujuan survei adalah untuk memberikan estimasi yang baik (dalam hal bias dan varians) dari beberapa karakteristik populasi melalui sampel. Dalam kerangka tersebut, kerepresentatifan dapat menjadi sarana untuk mencapai estimasi yang baik, bukan tujuan dari survei sampling itu sendiri.

Teori sampling probabilitas dan penggunaan informasi tambahan telah menunjukkan bahwa penggunaan *unequal probability* dapat meningkatkan akurasi estimasi. Dalam hal ini, sampel bisa sangat jauh dari menjadi miniatur populasi dan secara efisien mengestimasi karakteristik yang diinginkan. Neyman (1934) dalam makalahnya menekankan dua hal utama dari teori sampling probabilitas :

- i. Hanya metode *probability sampling* yang dapat secara teoritis divalidasi karena merupakan hasil percobaan random.
- ii. Hasil pada stratifikasi membuktikan bahwa sampel yang representatif tidak optimal dalam hal akurasi. Hal ini dapat dilihat sebagai kekalahan untuk *non-probability sampling*, serta untuk ide representatif sebagai tujuan di bidang survei sampling.

2.2 Konsep Dasar Sensus dan Survei

Sensus adalah cara pengumpulan data dimana semua unit (elemen) yang menjadi objek penelitian harus diteliti seluruhnya. Akibat seluruh unit (elemen) di dalam populasi diteliti di dalam sensus, maka dengan sendirinya diperlukan tenaga, waktu dan biaya yang sangat tinggi. Misalnya: Sensus Penduduk 2010 melakukan pendataan terhadap seluruh penduduk Indonesia. Di dalam metode survei sampel, penelitian hanya mengambil sebagian kecil dari unit-unit di dalam populasi untuk diteliti.

Tabel 2.1 Perbandingan Survei Sampel dan Sensus

Segi	Survei Sampel	Sensus
Tenaga	<ul style="list-style-type: none"> - Jumlah relatif sedikit - Dapat dipilih yang berkualitas 	<ul style="list-style-type: none"> - Jumlah sangat besar - Lebih sulit untuk memilih yang berkualitas seluruhnya
Waktu	<ul style="list-style-type: none"> - Lebih cepat 	<ul style="list-style-type: none"> - Lebih lama
Biaya	<ul style="list-style-type: none"> - Lebih murah 	<ul style="list-style-type: none"> - Lebih mahal
Pertanyaan dan kualitas data	<ul style="list-style-type: none"> - Biasanya kualitas data lebih baik - Pertanyaan yang lebih sulit bisa dipergunakan 	<ul style="list-style-type: none"> - Kualitas data kurang baik, hal ini akibat dari kualitas tenaga pengumpul - Pertanyaan sederhana
Penyajian Data	<ul style="list-style-type: none"> - Data tidak bisa disajikan sampai ke tingkat yang paling rendah 	<ul style="list-style-type: none"> - Data bisa disajikan sampai ke tingkat yang paling rendah, karena semua unit dalam populasi dikumpulkan
Kesalahan (<i>Error</i>)	<ul style="list-style-type: none"> - Adanya kesalahan sampel - Adanya kesalahan bukan dari sampel, tetapi relatif kecil 	<ul style="list-style-type: none"> - Tidak ada kesalahan sampel - Adanya kesalahan bukan dari sampel yang besar

Berdasarkan hasil penelitian sampel tersebut digunakan untuk mengestimasi nilai karakteristik populasi yang diteliti. Akibat hanya sebagian unit dalam populasi yang diteliti, maka jelas bahwa survei sampel akan lebih menghemat tenaga, waktu dan biaya dibandingkan dengan sensus. Beberapa hal yang menyebabkan survei sampel dilakukan di dalam penelitian (proses pengumpulan data) adalah:

- a. Populasinya tidak terbatas atau sangat besar.
- b. Terbatasnya biaya, tenaga dan waktu.
- c. Penelitian bersifat *destruktif* (merusak).
- d. Pengaturan manajemen pengumpulan data lebih terkendali.

2.3 Jenis Pengambilan Sampel

2.3.1 Sampling Tidak Berpeluang (*Non-Probability Sampling*)

Prosedur pengambilan sampel tidak berpeluang tergantung pada kebijakan dan pengalaman, tanpa memperhatikan kaidah-kaidah *probability*. Bias dan *sampling error* pengambilan sampel ini tidak dapat ditentukan berdasarkan sampel yang terpilih, sehingga kurang dapat dipertanggungjawabkan untuk analisis secara statistik. Adapun jenis-jenis pengambilan sampel tidak berpeluang yaitu:

- a. *Convenience sampling* yaitu pengambilan sampel yang semata-mata hanya mempertimbangkan kemudahan saja, oleh karena itu pengambilan sampel dengan cara ini tidak mewakili populasi dan hanya cocok untuk penelitian yang sifatnya *eksploratif* atau untuk *pilot study*. Misalnya untuk mempermudah penelitian, peneliti mengambil lima lembaga yang terdekat dengan rumahnya padahal belum tentu lembaganya memenuhi kriteria objek penelitian.
- b. *Purposive sampling* yaitu pengambilan sampel semata-mata menurut kriteria pemikiran dan pengetahuan pengambil sampel. Sampel yang terpilih sangat dipengaruhi sekali oleh pemahaman pengambil sampel terhadap karakteristik populasi. Metode ini sering digunakan dalam survei dengan jumlah unit sampel kecil. Misalnya : peneliti ingin memutuskan untuk menarik sampel satu kota yang mewakili populasi yang mencakup

seluruh kota. Ketika menggunakan metode ini, peneliti harus yakin bahwa sampel yang dipilih benar-benar mewakili dari seluruh populasi.

- c. *Quota sampling* yaitu pengambilan sampel dimana jumlah sampel telah ditentukan terlebih dahulu. Pengambil sampel tinggal memilih sampai jumlah tersebut dan biasanya tanpa kerangka sampel. Pengambilan sampel semacam ini sering digunakan dalam survei pendapat masyarakat. Misalnya : Survei kepuasan masyarakat DKI Jakarta terhadap pelayanan Bus Transjakarta.
- d. *Snowball sampling* yaitu pengambilan sampel yang dipakai ketika peneliti tidak banyak tahu tentang populasi penelitiannya sehingga dari beberapa sampel yang diambil dan diketahuinya, ia mengambil sampel lain dengan penjelasan dari sampel yang dikenalnya.

2.3.2 Sampling Berpeluang (*Probability Sampling*)

Prosedur pengambilan sampel berpeluang memperhatikan kaidah-kaidah *probability*, sehingga bias dan *sampling error* pengambilan sampel ini dapat ditentukan berdasarkan sampel yang terpilih. Hasil survei sampel hanya bisa untuk menduga nilai populasinya (parameter). Nilai penduga tersebut hanya mempunyai kemungkinan (*probability*) yang kecil untuk bisa sama dengan nilai populasinya.

Karena hanya sebagian yang akan dipilih dari unit yang ada dalam populasi, terdapat banyak pilihan kumpulan unit yang dapat diambil. Tiap kumpulan unit yang mungkin akan terambil sebagai sampel menghasilkan nilai pendugaan yang berbeda. Bila nilai-nilai unit di dalam populasi sama atau relatif hampir sama (homogen), dapat dikatakan bahwa hasil dugaan dari survei sampel adalah sama dengan nilai populasinya.

Untuk mengetahui bagaimana melakukan pengambilan sampel sehingga dapat memperkirakan tingkat kecermatannya, cara yang dapat digunakan adalah dengan menggunakan hukum-hukum peluang (acak) untuk penarikan unit ke dalam sampel. Cara ini dinamakan metode penarikan sampel berpeluang atau sering disingkat metode penarikan sampel. Pada metode ini setiap unit di dalam

populasi mempunyai peluang tertentu untuk terpilih sebagai anggota sampel. Jadi setiap anggota sampel sudah ditentukan nilai peluang untuk dapat terpilih.

Ada beberapa macam pemilihan sampel berpeluang antara lain :

a. Sampel Acak Sederhana (*Simple Random Sampling*)

Suatu sampel dinamakan sampel acak sederhana (*simple random sampling*) bila setiap unit dalam populasi diberi peluang sama untuk terpilih. Metode ini merupakan metode yang cukup mudah dan biasa digunakan pada populasi yang memuat karakteristik unit (unit) bersifat relatif homogen.

b. Sistematis Sampling (*Systematic Sampling*)

Suatu metode pengambilan sampel secara acak sistematis dengan interval (jarak) tertentu dari suatu kerangka sampel yang telah diurutkan.

c. Sampel Acak Berlapis (*Stratified Random Sampling*)

Sampel Acak Berlapis merupakan metode pemilihan sampel dimana berdasarkan suatu informasi (data) unit-unit di dalam populasi dikelompokkan-kelompokkan. Proses pembentukan kelompok-kelompok ini dinamakan stratifikasi. Diusahakan nilai-nilai unit di dalam suatu kelompok cukup homogen, sedangkan antar lapisan heterogen. Kelompok-kelompok semacam ini dinamakan lapisan (strata). Kemudian dari setiap lapisan yang dibentuk, dipilih sejumlah sampel secara random.

d. Sampel Acak Berkelompok (*Cluster Sampling*)

Prosedur sampling di mana unit terkecil dalam populasi tidak teridentifikasi secara lengkap hanya kelompok-kelompok dari unit-unit tersebut yang dapat diidentifikasi secara lengkap, di mana kelompok-kelompok itu disebut *cluster*. Pemilihan dilakukan terhadap *cluster-cluster*, bukan lagi terhadap unit-unit analisa terkecil. *Cluster-cluster* yang terpilih ke dalam sampel inilah yang selanjutnya menentukan semua unit-unit yang akan diselidiki.

Jadi, sampling berpeluang dapat dibedakan menjadi dua, yaitu: sampling dengan peluang sama (*equal probability*) dan sampling dengan peluang tidak sama (*unequal probability*).

Selain dibedakan menjadi sampling berpeluang dan tidak berpeluang, teknik sampling dapat pula dibedakan menjadi:

- a. Cara Tradisional: pengambilan sampel tanpa memperhatikan informasi sebelumnya dapat dilakukan dengan dua cara:
 - i. Sampling dengan teknik pengembalian: elemen-elemen dapat diambil dan dimasukkan kembali dalam sampel lebih dari sekali.
 - ii. Sampling dengan teknik tanpa pengembalian: elemen-elemen tidak dapat diambil atau dimasukkan kembali dalam sampel lebih dari sekali.
- b. Pendekatan Bayesian: informasi tentang parameter-parameter populasi sebelumnya sudah diketahui termasuk biaya dan probabilitas kesalahan, lebih bersifat teoritis, kenyataannya sulit dipenuhi karena informasi-informasi populasi susah didapat maka lebih sering dilakukan cara tradisional.

2.4 Desain Sampling Survei Angkatan Kerja Nasional (Sakernas) 2011

Pemilihan sampel rumah tangga dirancang dengan penarikan sampel tiga tahap, dengan tahapan sebagai berikut:

- a) Tahap pertama: dari daftar wilcah SP2010 dipilih 30.000 wilayah cacah untuk Susenas secara *Probability Proportional to Size (pps)* dengan *size* jumlah rumah tangga SP2010. Kemudian 30.000 wilayah cacah tersebut dialokasikan sama ke dalam empat triwulan, masing-masing sebesar 7.500 wilcah. Dari 7.500 wilcah Susenas Triwulan I, dipilih 5.000 wilayah cacah secara sistematis untuk Sakernas 2011 Triwulan I dan akan digunakan lagi untuk Triwulan II, III, dan IV. Wilayah cacah yaitu suatu wilayah yang terdiri dari sekitar 300 rumah tangga.
- b) Tahap kedua: memilih dua blok sensus pada setiap wilayah cacah terpilih Susenas yang juga terpilih Sakernas secara *pps* sistematis dengan *size* jumlah rumah tangga SP2010-C1. Selanjutnya blok-blok sensus terpilih dialokasikan secara acak untuk Susenas dan Sakernas. Blok-blok sensus terpilih Sakernas ini digunakan untuk estimasi provinsi dan dibagi ke dalam 4 paket sampel.

- c) Tahap ketiga: memilih 10 rumah tangga secara sistematis berdasarkan hasil pemutakhiran rumah tangga SP2010-C1.

Khusus untuk Sakernas Triwulan III, yang diperuntukkan untuk estimasi kabupaten, diperlukan tambahan sampel blok sensus. Dari 15.000 sampel wilcah terpilih Susenas Triwulan II dan III masing-masing dipilih 2 blok sensus, satu untuk keperluan Susenas dan yang lainnya untuk Sakernas. Blok sensus untuk Sakernas yang terpilih dari *PSU* Susenas Triwulan II dan III ini selanjutnya digunakan sebagai sampel blok sensus komplemen yang merupakan tambahan sampel yang apabila digabungkan dengan blok sensus estimasi provinsi (Sakernas Triwulan III) dapat digunakan untuk estimasi kabupaten (BPS, 2012). Tabel 2.2 menunjukkan rencana penarikan sampel Sakernas 2011 dalam bentuk tabel.

Tabel 2.2 Desain Penarikan Sampel Sakernas 2011

Tahap	Banyak unit pada strata t		Metode pemilihan sampel	Peluang sampel terpilih	Fraksi sampel	<i>Weight</i>
	Populasi	Sampel				
1	N_t	n_t	pps-sistematis	$\frac{w_{ti}}{W_t}$	$\frac{n_t \times w_{ti}}{W_t}$	$\frac{W_t}{n_t \times w_{ti}}$
2	M_{ti}	m_{ti}	pps-sistematis	$\frac{z_{tij}}{Z_{ti}}$	$\frac{m_{ti} \times z_{tij}}{Z_{ti}}$	$\frac{Z_{ti}}{m_{ti} \times z_{tij}}$
3	L_{tij}	l_{tij}	sistematis	$\frac{1}{L_{tij}}$	$\frac{l_{tij}}{L_{tij}}$	$\frac{L_{tij}}{l_{tij}}$

Berdasarkan Tabel 2.2, dapat diketahui rumus estimasi total dari suatu karakteristik/ variabel sebagai berikut:

$$\hat{X} = \sum_{t=1}^T \frac{W_t}{n_t} \sum_{i=1}^{n_t} \frac{Z_{ti}}{w_{ti} \times m_{ti}} \sum_{j=1}^{m_{ti}} \frac{L_{tij}}{z_{tij} \times l_{tij}} \sum_{k=1}^{l_{tij}} x_{tijk} \quad (2.1)$$

Keterangan:

- \hat{X} = estimasi total dari suatu variabel
 t = 1, ..., T ; strata t dari 1 sampai dengan T
 n_t = jumlah sampel wilayah cacah pada strata ke- t
 W_t = jumlah total rumah tangga pada strata ke- t

w_{ti}	=	jumlah rumah tangga pada strata ke- t dan sampel wilayah cacah ke- i
m_{ti}	=	jumlah sampel blok sensus pada strata ke- t dan sampel wilayah cacah ke- i
Z_{ti}	=	jumlah total rumah tangga pada strata ke- t dan sampel wilayah cacah ke- i
z_{tij}	=	jumlah rumah tangga pada strata ke- t , sampel wilayah cacah ke- i , dan sampel blok sensus ke- j
l_{tij}	=	jumlah sampel rumah tangga pada strata ke- t , sampel wilayah cacah ke- i , dan sampel blok sensus ke- j
L_{tij}	=	jumlah total rumah tangga pada strata ke- t , sampel wilayah cacah ke- i , dan sampel blok sensus ke- j
x_{tijk}	=	nilai karakteristik dari rumah tangga ke- k pada strata ke- t , sampel wilayah cacah ke- i , dan sampel blok sensus ke- j

Estimasi total variabel dari desain sampling BPS pada penelitian ini menggunakan penghitungan sebagai berikut:

$$\hat{X} = \sum_t \frac{W_t}{n_t} \sum_k \frac{x_{tk}}{w_{tk}} \quad (2.2)$$

dengan:

\hat{X}	=	estimasi jumlah nilai karakteristik
n_t	=	jumlah sampel pada strata ke- t
W_t	=	jumlah total rumah tangga pada strata ke- t
w_{tk}	=	jumlah rumah tangga sampel pada strata ke- t dan unit sampel ke- k
x_{tk}	=	nilai karakteristik dari unit sampel/ blok sensus ke- k pada strata ke- t

Persamaan (2.2) merupakan penyederhanaan/ pendekatan dari persamaan (2.1) dengan estimasi desain sampling BPS. Hal ini dilakukan dengan pertimbangan sebagai berikut:

- Unit sampel pertama pada Sakernas yang blok sampelnya digunakan pada penelitian ini adalah rumah tangga. Namun, pada penelitian ini unit sampel yang digunakan adalah blok sensus. Jadi, informasi tiap blok sensus sudah didapatkan secara langsung dari data SP2010 sehingga tidak memerlukan

estimasi pada tingkat blok sensus dari rumah tangga terpilih sebagaimana estimasi dari survei Sakernas.

- b. Tidak tersedianya data wilayah cacah yang bersesuaian dengan blok sensus pada *raw* data SP2010. Selain itu, untuk estimasi tingkat provinsi terdapat sampel sejumlah 116 blok sensus di Provinsi Kalimantan Tengah. Bila dilihat dari tiga tahapan di Tabel 2.2, maka pada tahap pertama diperoleh 116 wilayah cacah di Kalimantan Tengah. Pada tahap kedua diperoleh 116 blok sensus untuk susenas dan 116 blok sensus untuk Sakernas. Jika hanya fokus pada Sakernas, maka sebenarnya pada tahap kedua hanya memilih satu blok sensus dari tiap wilayah cacah yang telah diperoleh pada tahap pertama.

2.5 Estimasi Rasio dan Estimasi Varians Rasio Dua Nilai Karakteristik

Bila terdapat dua variabel Y dan X dan memiliki hubungan rasio sebagai berikut:

$$\hat{R} = \frac{\hat{Y}}{\hat{X}} = \frac{\sum_{t=1}^T \sum_{i=1}^{n_t} \sum_{k=1}^{m_i} P_{tik} y_{tik}}{\sum_{t=1}^T \sum_{i=1}^{n_t} \sum_{k=1}^{m_i} P_{tik} x_{tik}} = \frac{\sum_{t=1}^T \sum_{i=1}^{n_t} \hat{Y}_{ti}}{\sum_{t=1}^T \sum_{i=1}^{n_t} \hat{X}_{ti}} \quad (2.3)$$

Maka estimate varians \hat{R} dengan metode linierisasi Taylor (*Taylor Linearization*) yang merupakan salah satu metode untuk mengestimasi varians bagi total, rasio ataupun rata-rata adalah sebagai berikut:

$$\text{var}(\hat{R}) = \sum_{t=1}^T \frac{(1-f_t)}{\hat{X}_t^2} \left[\frac{n_t}{n_t-1} \left(\sum_{i=1}^{m_t} \hat{D}_{ti}^2 - \frac{\hat{D}_t^2}{n_t} \right) \right] \quad (2.4)$$

dengan:

P_{tik} = penimbang/ *weight* rumah tangga ke- k , blok sensus ke- i dalam strata t

$\hat{D}_{ti} = \hat{Y}_{ti} - \hat{R}_t \cdot \hat{X}_{ti}$, dan $\hat{D}_t = \hat{Y}_t - \hat{R}_t \cdot \hat{X}_t$

n_t = jumlah blok sensus terpilih dalam strata t ,

\hat{Y}_{ti} = estimasi jumlah nilai karakteristik Y dalam blok sensus i dalam strata t ,

\hat{X}_{ti} = estimasi jumlah nilai karakteristik X dalam blok sensus i dalam strata t ,

f_t = fraksi penarikan sampel blok sensus yang pada umumnya bernilai kecil, maka biasanya tidak diperhitungkan.

2.6 Karakteristik Estimator dari Survei dan Desain Sampling yang Baik

Menurut Levy dan Lemeshow (1999), estimator parameter populasi memiliki beberapa karakteristik sebagai akibat dari penggunaan sampling, yaitu:

- a. Bias $B(\hat{\theta})$ didefinisikan sebagai nilai selisih antara nilai harapan $E(\hat{\theta})$ dengan nilai sebenarnya (θ) , sehingga

$$B(\hat{\theta}) = E(\hat{\theta}) - \theta \quad (2.5)$$

Estimator dikatakan tidak bias jika nilai selisih tersebut sama dengan nol.

- b. *Mean Square Error* $MSE(\hat{\theta})$ didefinisikan sebagai rata-rata simpangan kuadrat estimator total populasi dengan parameter populasi sebenarnya dikalikan peluang π_i , sehingga

$$MSE(\hat{\theta}) = E[(\hat{\theta} - \theta)^2] \quad (2.6)$$

MSE berbeda dengan varians karena MSE merupakan rata-rata simpangan kuadrat terhadap parameter sebenarnya. Adapun varians merupakan rata-rata simpangan kuadrat terhadap rata-rata distribusi penarikan contoh.

Hubungan MSE dan varians adalah sebagai berikut:

$$MSE(\hat{\theta}) = Var(\hat{\theta}) + (B(\hat{\theta}))^2 \quad (2.7)$$

- c. Reliabilitas. Karakteristik terandal dari suatu estimator populasi dengan bagaimana suatu estimator (melalui pengulangan) dalam proses menghasilkan suatu nilai estimasi. Jika kita mengasumsikan tidak ada kesalahan pengukuran dalam suatu survey, maka reliabilitas dapat dinyatakan dalam konteks varians atau setara dengan *standar error*. Makin kecil *standar error*, maka semakin besar reliabilitasnya.
- d. Validitas. Karakteristik yang berhubungan dengan bagaimana nilai tengah suatu estimator dalam proses menghasilkan suatu estimasi berbeda dengan nilai parameter sebenarnya. Jika kita mengasumsikan tidak ada kesalahan pengukuran, validitas dapat dievaluasi dengan mengamati nilai bias dari estimatornya. Semakin kecil bias, validitas semakin besar.
- e. Akurasi dari suatu estimator berhubungan dengan sejauh mana rata-rata suatu nilai estimasi menyimpang dari nilai parameter yang diukur. Akurasi suatu estimator pada umumnya dievaluasi oleh nilai MSE-nya, atau setara dengan nilai akar pangkat MSE, yaitu *Root Mean Square Error* atau

RMSE. Semakin kecil nilai MSE, semakin besar nilai akurasi. Akurasi mencakup reliabilitas dan validitas.

Levy dan Lemeshow (1999) juga menyebutkan bahwa ada tiga kriteria desain sampling yang baik, yaitu memiliki tingkat akurasi estimator yang tinggi, biaya paling kecil, dan feasibilitas dalam melaksanakan desain sampling tersebut. Dalam penelitian ini, perbandingan antar desain sampling hanya akan menggunakan kriteria tingkat akurasi.

2.7 *Balanced Sampling*

Konsep *Balanced Sampling* mulai muncul dari pelaksanaan survei pada tahun 1929 di Italia. Pada saat itu, pimpinan Kantor Statistik Italia, Gini dan Galvani, memilih sampel sebesar 29 distrik dari 214 distrik yang ada. Pemilihan 29 distrik ini didasarkan pada kesesuaian dengan tujuh variabel kontrol yang datanya berasal dari sensus 1921. Sampel dipilih dengan melihat rata-rata sampel yang paling mendekati rata-rata populasi (Langel dan Tille, 2011).

Brewer dan Gregoire (2009) menuliskan bahwa hal yang dilakukan oleh Gini dan Galvani ini dikritik oleh Neyman (1934):

- a. Pertama, konsep Gini Galvani tidak menggunakan prinsip randomisasi sehingga tidak dapat menggunakan *Central Limit Theorem*. Selain itu, tidak dapat dibentuk *confident interval*.
- b. Kedua, kesulitan mendapatkan kebutuhan “*purposive*” (7 variabel) menyebabkan hanya mempertimbangkan 214 distrik tidak melihat 8.354 *commune* yang ada di Italia.
- c. Ketiga, model populasi yang digunakan ditunjukkan dengan tidak realistis dan tidak sesuai.

Yates (1960) dalam Langel dan Tille (2011) juga memberikan kritikan terhadap metode pemilihan *purposive* dari paper Gini dan Galvani pada tahun 1929. Yates menggarisbawahi pada fakta bahwa reabilitas dari hasil tidak dapat dinilai dengan metode *purposive*. Namun, Yates menekankan bahwa tidak ada kontradiksi antara *balanced sampling* dengan *random sampling*. Oleh karena itu, Yates mengusulkan sebuah metode memilih sebuah *random balanced sample*. Randomisasi dilakukan dengan memilih sampel random di awal dan kemudian

memilih unit berikutnya. Jika unit baru meningkatkan keseimbangan, maka dijadikan sampel, tetapi jika unit baru tidak meningkatkan keseimbangan, maka tidak dijadikan sampel.

Royal dan Herson (1973) menulis mengenai tiga jenis *balanced sampling*. Selain *purposive* dan random sampling yang sedikit disinggung di atas, disebutkan pula *restricted randomization*. *Restricted randomization* didefinisikan sebagai sebuah proses pemilihan yang memberikan pendekatan *balanced sample* dengan tidak mengabaikan randomisasi. Ketiga metode ini tidak menjamin bahwa sebuah *balanced sample* juga akan *balanced* pada variabel yang tidak dimasukkan dalam proses pemilihan. Royal dan Herson (1973) dalam Berger dan Tille (2009) menegaskan pentingnya *balanced sample* adalah untuk melindungi inferensi dari *misspecified model*. Mereka menyebutnya *robustness*. *Robustness* ini didapatkan karena momen sampel dari variabel tambahan berada di sekitar momen populasi variabel tambahan (Lohr, 2009).

Langel dan Tille (2011) menjelaskan pada awalnya *balanced sampling* diklasifikasikan sebagai penarikan sampel *purposive*, sebagaimana yang diaplikasikan oleh Gini dan Galvani. Akan tetapi, metode *balanced sampling* kemudian berkembang dengan tidak berkontradiksi dengan kerangka *random sampling*. Metode yang modern memungkinkan pemilihan *balanced sample* dan pada saat yang sama tetap konsisten pada gagasan kerandoman. Deville dkk (1988) mengusulkan sebuah metode pemilihan *balanced sample* dengan beberapa variabel dengan peluang terpilih yang sama. Metode ini didasarkan pada konstruksi grup yang dibangun sesuai rata-rata dari variabel tambahan. Deville dan Tille (1998) mengusulkan delapan metode baru, salah satu di antaranya *splitting method*. Metode ini merupakan cara untuk menyajikan dengan lebih sederhana beberapa metode sampling yang telah ada. Metode ini kemudian dikembangkan lagi oleh Deville dan Tille (2004) menjadi *cube method*. Penelitian lain mengenai *balanced sampling* di dalam Langel dan Tille (2011) dilakukan oleh Deville (1992), Deville dan Tille (2004), Hedayat dan Majumdar (1995), Nedyalkova dan Tille (2009), serta Tille dan Favre (2004), (2005).

2.7.1 Definisi *Balanced Sampling*

Sebuah desain sampling disebut *balanced* jika estimator dari sebuah set variabel tambahan sama dengan total variabel tambahan dari populasi yang diketahui. *Balanced sampling* dapat dilihat sebagai metode kalibrasi/ pencocokan yang melekat pada desain sampling (Berger dan Tille, 2009).

Jika nilai rata-rata dari beberapa variabel diketahui untuk setiap unit dari seluruh populasi, maka dimungkinkan untuk memperoleh sampel yang rata-ratanya *equal*/ sama dengan rata-rata populasi. Sampel seperti ini ekuivalen dengan sampel random, pada kasus ini dinamakan *balanced sample* (Yates, 1949).

Langel dan Tille (2011) mendefinisikan *balanced sampling* dengan menggunakan konsep sampel yang representatif sebagai miniatur dari sebuah populasi, yaitu dalam paradigma pemilihan *purposive* dan dalam kerangka *random sampling*. Sebuah sampel dikatakan *balanced* jika estimator alami dari total variabel tambahan X akan bersesuaian (atau mendekati) dengan total dari variabel pada populasi yang diketahui. Estimator di sini adalah sebuah estimator dimana bobot satuannya tidak berubah dari satu sampel ke sampel yang lain.

Deville dan Tille (2004) mendefinisikan desain *balanced sampling* dengan *property* bahwa estimator Horvits-Thompson dari total populasi sebuah set variabel tambahan sama dengan total variabel yang diketahui. Tille (2006) memberikan definisi *balanced sampling* secara matematis sebagaimana berikut:

$$Y = \sum_{k \in U} y_k ,$$

dimana Y = total nilai dari variabel pada U

U = set unit di dalam populasi

k = unit dari populasi

y = variabel yang diminati

y_k = nilai dari variabel y untuk unit k

Misalkan vektor nilai dari p variabel tambahan yang diketahui untuk semua unit populasi adalah sebagai berikut:

$$\mathbf{x}_k = (x_{k1} \dots x_{kj} \dots x_{kp})'$$

dimana p = jumlah variabel *balancing*

\mathbf{x}_k = vektor nilai dari p variabel tambahan pada unit k dari populasi

x_j = variabel tambahan ke- j atau variabel *balancing*

x_{kj} = nilai dari variabel x_j untuk unit k

Vektor total populasi dari variabel *balancing* adalah

$$\mathbf{X} = \sum_{k \in U} \mathbf{x}_k,$$

dan dapat diestimasi dengan

$$\hat{\mathbf{X}}_{HT} = \sum_{k \in U} \frac{\mathbf{x}_k S_k}{\pi_k}.$$

dengan $S_k = \begin{cases} 1 & \text{jika } k \text{ merupakan sampel} \\ 0 & \text{jika } k \text{ bukan sampel} \end{cases}$

π_k = peluang terpilih dari unit k

$\hat{\mathbf{X}}_{HT}$ = estimator Horvits-Thompson dari vektor total populasi dari variabel *balancing*

Selanjutnya akan dibangun sebuah desain *balanced sampling*, sesuai Definisi 1.

Definisi 1. Sebuah desain sampling dengan peluang memilih sampel s , $p(s)$, dikatakan *balanced* dengan memperhatikan variabel tambahan x_1, \dots, x_p , jika dan hanya jika memenuhi persamaan *balancing* berikut:

$$\hat{\mathbf{X}}_{HT} = \mathbf{X} \quad (2.8)$$

yang dapat ditulis

$$\sum_{k \in U} \frac{x_{kj} S_k}{\pi_k} = \sum_{k \in U} x_{kj},$$

untuk semua $s \in S$ sehingga $p(s) > 0$ dan untuk semua $j = 1, \dots, p$; dengan perkataan lain varians estimator Horvits-Thompson

$$\text{var}(\hat{\mathbf{X}}_{HT}) = 0.$$

Dengan demikian, *balanced sampling* dapat dipandang sebagai batasan pada himpunan sampel Q dan hanya sampel yang memenuhi persamaan *balancing* yang memiliki peluang positif. Himpunan sampel Q ini dapat dituliskan sebagai berikut:

$$Q = \left\{ s \in S \mid \sum_{k \in U} \frac{\mathbf{x}_k s_k}{\pi_k} = \mathbf{X} \right\}$$

2.7.2 Tujuan *Balanced Sampling*

Opsomer (2011) menjelaskan bahwa setidaknya ada dua tujuan *balanced sampling* yaitu memperoleh representasi struktur populasi pada sampel dan meningkatkan efisiensi dari estimator survey. Saat ini, para statistisi mencapai kedua tujuan ini dengan menggunakan stratifikasi dan sistematis. Stratifikasi mencapai *balance* dengan membagi sampel sesuai alokasi stratanya dan meningkatkan efisiensi estimator dengan meniadakan komponen varians yang disebabkan oleh perbedaan di antara strata. Sampling sistematis juga digunakan untuk mencapai tujuan ini. Hal ini didapat dengan komposisi sampel yang sesuai komposisi populasi sesuai urutan dalam suatu variabel dan variabel lain yang berkorelasi dengan variabel pengurut. Efisiensi sampel didapatkan pada saat momen sampel akan berada di sekitar dari momen populasi. Kedua metode ini tidak fleksibel. Stratifikasi membagi populasi ke dalam sel-sel yang didefinisikan oleh variabel stratifikasi, yang mungkin menyebabkan timbulnya beberapa sel kecil. Sampling sistematis merupakan bentuk sampling yang sangat terbatas dengan sejumlah fleksibilitas di dalam konstruksi sampel. *Balanced sampling* dapat dilihat sebagai generalisasi dari stratifikasi, tetapi dapat menggunakan variabel apapun, baik variabel kategorik, variabel kontinu, ataupun kombinasi dari keduanya. Hal ini memberikan fleksibilitas dalam konstruksi sampel.

2.7.3 *Cube Method*

Deville dan Tille (2004) mengusulkan *cube method* sebagai sebuah metode umum yang memungkinkan pemilihan *balanced sample* dimana estimasi Horvitz-Thompson untuk variabel tambahan adalah sama, atau mendekati sama, dengan parameter variabel tambahan populasinya. Metode ini dapat digunakan untuk sejumlah besar variabel *balancing* bersifat kualitatif dan kuantitatif. Metode ini juga memungkinkan akan menjamin peluang terpilihnya suatu unit adalah sama atau tidak sama dan memungkinkan untuk mengetahui seberapa akurat sebuah

sampel dapat disetimbangkan. Maksud dari *cube method* adalah sebagai metode umum yang dapat digunakan untuk menentukan peluang suatu unit dapat terpilih adalah sama atau tidak sama. Selain itu, metode ini juga dapat dikombinasikan dengan metode sampling yang telah ada, seperti *cluster sampling*, *multistage sampling*, ataupun *stratified sampling*.

Berger dan Tille (2009) menyebutkan bahwa *cube method* telah digunakan untuk memilih rotasi grup dari sensus di Perancis dan pemilihan master sampel Perancis. Deville dan Tille (2004) menjelaskan nama *cube method* berdasarkan representasi geometri dari desain sampling sebagaimana Gambar 2.1 (a). Sampel yang mungkin sebesar 2^N berkorespondensi dengan 2^N dalam vektor R^N . Setiap vektor sampel s adalah sebuah *vertex* (sudut) dari sebuah N -cube, dan jumlah sampel yang mungkin adalah jumlah *vertice* (beberapa sudut) dari N -cube. Sebuah desain sampling dengan peluang terpilih $\pi_k (k \in U)$ terdiri dari penetapan sebuah peluang $p(s)$ pada tiap *vertex* (sudut/ sampel) dari N -cube sehingga

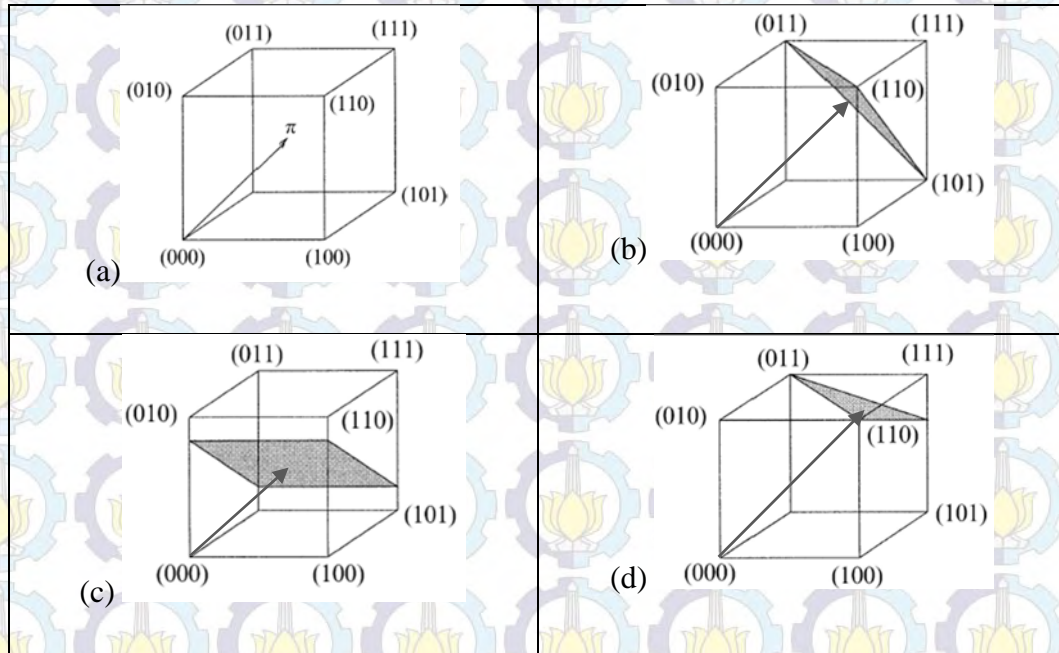
$$E(s) = \sum_{s \in S} p(s) = \pi,$$

di mana $\pi = (\pi_k)$ adalah vektor peluang terpilih.

Berger dan Tille (2009) menjelaskan bahwa sebuah sampel dapat ditulis sebagai vektor $s = (s_1, \dots, s_N) \in R^N$ dari indikator variabel s_k yang mengambil nilai 1 untuk unit terpilih dan 0 untuk tidak terpilih. Algoritma samplingnya dapat diilustrasikan sebagai sebuah cara random untuk mencapai *vertex* (sudut) dari N -cube berupa vektor π sehingga memenuhi persamaan (2.8).

Berikut gambaran tentang representasi geometri dari *cube method*. Misalkan terdapat populasi sebanyak $N = 3$ dengan jumlah kemungkinan sampel pada *cube method* adalah 2^N . Dengan demikian, populasi sebanyak 3 memiliki 2^3 kemungkinan sampel, yaitu 8 kemungkinan sampel. Kemungkinan sampel ini digambarkan sebagai set sudut dari sebuah kubus. Sudut (000) menunjukkan sampel kosong, yaitu tidak ada satu pun sampel yang dipilih. Sudut (100), (010), dan (001) menunjukkan kemungkinan terpilih satu sampel. Sampel tersebut dapat unit pertama, kedua, atau ketiga. Selanjutnya, sudut (110), (101), dan (011) menunjukkan kemungkinan terpilih sebanyak dua sampel. Sampel dapat merupakan unit pertama dan kedua, unit pertama dan ketiga, ataupun kedua dan

ketiga. Sudut terakhir adalah (111), yaitu terpilihnya ketiga unit yang ada (sensus). Gambar 2.1(a) menunjukkan representasi geometrik dari kemungkinan sampel dengan populasi sebanyak $N = 3$.



Gambar 2.1 (a) Representasi Geometrik dari Kemungkinan Sampel dari Populasi dengan $N=3$; (b) *Balanced Constrain* ukuran tetap dari jumlah sampel=2: sampel *balanced* ada; (c) *Balanced constrain* sedemikian sehingga sampel *balanced* tidak ada; (d) *Balanced Constrain* sedemikian sehingga sampel *balanced* ada yang dapat diwujudkan dalam sudut kubus dan ada yang tidak

2.7.4 Tahapan *Balanced Sampling* dengan *Cube Method*

Proses *Balanced Sampling* dengan *cube method* dilakukan melalui tiga tahapan, yaitu: *Flight Phase*, *Landing Phase*, dan *Quality Balancing*. Detail tahapan diberikan dalam Algoritma 2.1 dan 2.2.

1. *Flight Phase*

Berger dan Tille (2009) menyebutkan bahwa *flight phase* adalah sebuah *random walk* yang dimulai pada vektor peluang terpilih dan masih di dalam interseksi dari kubus dan *constrain subspace* Q . *Random walk* ini berhenti pada *vertex* (sudut) dari interseksi dari *cube* dan *constrain subspace*. Tille (2011) menjelaskan bahwa pada *cube method*, jika n

jumlah sampel telah ditetapkan, *flight phase* akan secara random mentransformasi vektor peluang terpilih menjadi vektor yang berisi 0 dan 1. Pada setiap step, vektor peluang terpilih ditransformasi secara random, tetapi penjumlahan dari unit di dalam vektor akan sama dengan jumlah sampel yang telah ditentukan.

Berikut gambaran proses transformasi peluang terpilih di *flight phase* yang menghasilkan penjumlahan unit vektor peluang terpilih sama dengan jumlah sampel. Misalkan vektor peluang terpilih $\pi = (0,5 \ 0,5 \ 0,5 \ 0,5)'$ sehingga $n=2$. Kemudian vektor peluang terpilih mengalami transformasi dalam *flight phase* (Tille, 2011). Penjumlahan anggota vektor hasil transformasi sama dengan jumlah sampel yang telah ditetapkan. Hasil seperti ini sesuai dengan Gambar 2.1 (b). Gambar 2.1 (b) menunjukkan bahwa *hyperplane* memotong sudut dari kubus. Kondisi ini menunjukkan penjumlahan peluang terpilih berupa angka bulat dan persamaan *balancing* terpenuhi, dengan demikian maka sampel *balanced* ada.

$$\pi = \begin{pmatrix} 0,5 \\ 0,5 \\ 0,5 \\ 0,5 \end{pmatrix} \rightarrow \begin{pmatrix} 0,66 \\ 0,66 \\ 0,66 \\ 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 \\ 0,5 \\ 0,5 \\ 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

Namun, ada kalanya tidak semua unit dalam vektor peluang terpilih tertransformasi menjadi 0 atau 1 sehingga penjumlahan unit tidak sama dengan n jumlah sampel yang ditentukan. Berikut gambaran proses transformasi peluang terpilih di *flight phase* yang menghasilkan penjumlahan unit vektor peluang terpilih yang tidak sama dengan jumlah sampel. Misalkan vektor peluang terpilih $\pi = (0,5 \ 0,5 \ 0,5 \ 0,5 \ 0,5)'$ dan $n=3$, pada akhir *flight phase* dihasilkan vektor sebagai berikut:

$$\pi = \begin{pmatrix} 0,5 \\ 0,5 \\ 0,5 \\ 0,5 \\ 0,5 \end{pmatrix} \rightarrow \begin{pmatrix} 0,625 \\ 0 \\ 0,625 \\ 0,625 \\ 0,625 \end{pmatrix} \rightarrow \begin{pmatrix} 0,5 \\ 0 \\ 1 \\ 1 \\ 0,5 \end{pmatrix} \rightarrow \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0,25 \end{pmatrix} \rightarrow \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \rightarrow \pi^*$$

Hasil di atas sesuai dengan Gambar 2.1 (d). Gambar 2.1 (d) menunjukkan bahwa *hyperplane* memotong dua sudut dari kubus, tetapi salah satunya tidak. Kondisi ini menunjukkan persamaan *balancing* kadang terpenuhi. Kondisi lain yang mungkin terjadi dari hasil *flight phase* sebagaimana Gambar 2.1 (c). Gambar 2.1 (c) menunjukkan bahwa *hyperplane* tidak memotong sudut dari kubus. Kondisi ini menunjukkan persamaan *balancing* tidak terpenuhi dengan pasti. Tidak ada satupun sampel yang terpilih.

Chauvet dan Tille (2006) dalam Tille (2006) mengusulkan sebuah algoritma *fast flight phase* dimana perhitungan waktu meningkat secara linier dengan ukuran populasi sesuai Algoritma 2.1.

Algoritma 2.1 Algoritma *Fast Flight Phase*

Step 1. Inisialisasi

- a) Unit dengan peluang terpilih mendekati 0 atau 1 dikeluarkan dari populasi sebelum mengaplikasikan algoritma sehingga semua unit yang masih ada memiliki peluang $0 < \pi_k < 1$.
- b) Peluang terpilih diisikan ke dalam vektor π .
- c) Vektor ψ terdiri dari $p + 1$ pertama dari π .
- d) Dibuat sebuah vektor ranking $\mathbf{r} = (1 \ 2 \ \dots \ p \ p+1)'$.
- e) Matriks \mathbf{A} merupakan matrik penimbang
- f) Matriks \mathbf{B} terdiri dari $p + 1$ pertama kolom \mathbf{A}
- g) Inisial $k = p + 2$

Step 2. Proses Sampling

- a) Sebuah vektor \mathbf{u} diambil dari kernel \mathbf{B} .

b) Hitung λ_1^* dan λ_2^* , nilai terbesar dari λ_1 dan λ_2 sehingga $0 \leq \psi + \lambda_1 \mathbf{u} \leq 1$ dan $0 \leq \psi + \lambda_2 \mathbf{u} \leq 1$. Catat bahwa $\lambda_1^* > 0$ dan $\lambda_2^* > 0$.

c) Pilih $\psi = \begin{cases} \psi + \lambda_1 \mathbf{u} & \text{dengan peluang } q \\ \psi + \lambda_2 \mathbf{u} & \text{dengan peluang } 1 - q \end{cases}$
di mana $q = \lambda_2^* / (\lambda_1^* + \lambda_2^*)$.

d) Unit yang berkorespondensi dengan angka integer $\psi(i)$ dihapus dari \mathbf{B} dan diganti dengan unit baru.

FOR i=1 to p+1, DO

IF $\psi(i) = 0$ or $\psi(i) = 1$ THEN

IF $k \leq N$ THEN $\pi(\mathbf{r}(i)) = \psi(i)$;

$\mathbf{r}(i) = k$;

$\psi(i) = \pi(k)$;

FOR j = 1 to p, DO $\mathbf{B}(j, i) = \mathbf{A}(j, k)$;

ENDFOR;

$k = k + 1$

ELSE GOTO STEP 3.;

ENDIF;

ENDIF;

ENDFOR.

e) Kembali ke Step 2.a)

Step 3. Akhir dari bagian *flight phase*

For i = 1 to p+1, DO $\pi(\mathbf{r}(i)) = \psi(i)$ ENDFOR

Keterangan:

π = vektor peluang terpilih

ψ = vektor yang terdiri dari p + 1 pertama dari π

- \mathbf{r} = vektor ranking $(1, 2, \dots, p, p+1)'$
 \mathbf{B} = matriks yang terdiri dari $p+1$ pertama kolom \mathbf{A}
 \mathbf{A} = matriks $p \times N$ dari vektor penimbang $(\bar{\mathbf{x}}_1 \dots \bar{\mathbf{x}}_k \dots \bar{\mathbf{x}}_N)$; $\bar{\mathbf{x}}_k = \mathbf{x}_k / \pi_k$
 \mathbf{u} = vektor dari kernel matriks \mathbf{B} , dimana $\{\mathbf{u} | \mathbf{B}\mathbf{u} = \mathbf{0}\}$
 ψ = parameter desain eksponensial
 i = indeks pada vektor ψ dan \mathbf{r} serta indeks pada kolom matriks \mathbf{B} ;
 $i=1, \dots, p+1$
 j = indeks pada baris matriks \mathbf{B} matriks \mathbf{A} ; $j=1, \dots, p$
 p = jumlah variabel *balancing* yang digunakan

2. *Landing Phase*

Landing phase dimulai pada akhir dari *flight phase*. Jika sebuah sampel tidak didapatkan pada akhir *flight phase*, maka sebuah sampel dipilih mendekati kemungkinan *constrain subspace*. *Flight phase* berhenti pada sebuah *vertex* yang dinotasikan π^* dari interseksi di antara *cube* dengan Q . Hal ini dapat ditunjukkan bahwa

$$\text{card } U^* = \text{card} \{k \in U | 0 < \pi_k^* < 1\} = q \leq p,$$

yang berarti bahwa jumlah noninteger elemen dari π^* lebih kecil atau sama dengan jumlah variabel *balancing*. Tujuan dari *landing phase* adalah untuk menemukan sebuah sampel random s sehingga $E(s | \pi^*) = \pi^*$ dan yang sangat mendekati *balanced*.

Dua cara dapat digunakan untuk memilih sampel, yaitu:

- Metode pertama terdiri dari perhitungan semua sampel yang konsisten dengan π^* , sebuah sampel s menjadi konsisten jika $s_k = \pi_k^*$ ketika π_k^* adalah sebuah integer. Kemudian fungsi biaya $C(s)$ dilekatkan pada tiap sampel. Deville dan Tille (2004) mengusulkan beberapa $C(s)$. Dengan menggunakan metode *linier programming* memungkinkan untuk menemukan desain sampling pada sampel yang konsisten yang memenuhi peluang terpilihnya π^* dan yang meminimalkan rata-rata biaya. Akhirnya, sebuah sampel dipilih secara random dengan

mengikuti desain sampling ini. Metode ini dapat digunakan dengan jumlah variabel *balancing* yang lebih kecil dari 15 karena ini memerlukan perhitungan 2^{15} sampel. Hal ini sesuai Algoritma 2.2.

Algoritma 2.2 Landing Phase dengan Algoritma Enumerative

- Selesaikan $\min_{p(s)} \sum_{s \in S} p^*(s)C(s)$ dengan batasan $\sum_{s \in S} p^*(s) = 1$,
 $\sum_{s \in S} sp^*(s) = \pi^*$, dan $0 \leq p^*(s) \leq 1$ untuk semua $s \in S$
- Bentuk list $\{s_1, s_2, \dots, s_v, \dots, s_V\}$ dari semua kemungkinan sampel dengan semua peluangnya
- Bangkitkan variabel *random* u yang berdistribusi uniform (0,1)
- Pilih sampel s_j sehingga $\sum_{c=1}^{v-1} p^*(s_c) \leq u \leq \sum_{c=1}^v p^*(s_c)$

Keterangan:

S = semua kemungkinan sampel sebanyak 2^q

$p^*(s)$ = probability memilih sampel pada *landing phase*

$C(s)$ = fungsi biaya yang melekat pada sampel

V = kemungkinan sampel pada *landing phase*

c = indeks kemungkinan sampel pada *landing phase*

$\sum_{c=1}^{v-1} p^*(s_c)$ = kumulatif peluang sampel terpilih sebelum kemungkinan sampel ke- c

$\sum_{c=1}^v p^*(s_c)$ = kumulatif peluang sampel terpilih sampai dengan kemungkinan sampel ke- c

- Metode kedua dapat digunakan ketika jumlah variabel tambahan terlalu besar untuk solusi dari algoritma yang sederhana. Pada akhir dari *flight phase*, sebuah variabel tambahan dapat dikeluarkan.

Selanjutnya, satu dapat kembali pada *flight phase* hingga tidak mungkin dipindah dengan batasan *subspace*-nya.

3. *Quality Balancing*

Sebelumnya telah disebutkan bahwa Deville dan Tille (2004) mendefinisikan desain *balanced sampling* dengan *property* bahwa estimator Horvitz-Thompson dari total populasi sebuah set variabel tambahan adalah *equal*/ sama dengan total variabel yang diketahui. Perbandingan estimator dengan nilai sebenarnya berada pada tahapan ini. *Quality balancing* dilihat dengan *Relative Deviation* (RD) sesuai dengan persamaan (2.9).

$$RD = 100 \times \frac{\hat{X}_{jHT} - X_j}{X_j}, \quad (2.9)$$

dengan

\hat{X}_{jHT} = Estimator Horvitz-Thompson dari X_j

X_j = Jumlah nilai sebenarnya dari variabel tambahan ke- j

Tille (2006) menggunakan estimator Horvitz-Thompson dan varians dari estimator Horvitz-Thompson untuk *balanced sampling*. Jika X adalah variabel yang digunakan dalam proses penarikan *balanced sampling* dan Y merupakan variabel yang diteliti di dalam survey, maka estimator Y adalah sebagai berikut:

$$\hat{X}_{HT} = \sum_{k \in U} \frac{x_k S_k}{\pi_k} \quad (2.10)$$

dimana S_k = jumlah berapa kali unit k termasuk dalam sampel S

2.8 Metode Bayesian

Ada beberapa cara dalam mengestimasi parameter. Salah satu yang mulai populer adalah dengan pendekatan Bayesian. Berbeda dengan pendekatan lain yang hanya menggunakan data sampel, Bayesian juga menggunakan prior (pengalaman/ distribusi awal) dari data tersebut.

Pada pendekatan klasik, parameter dianggap bernilai tetap, sedangkan pada Bayesian parameter adalah variabel yang memiliki distribusi, yaitu distribusi

prior. Dengan distribusi prior ini, dapat ditentukan distribusi posterior. Kemudian dengan menggunakan konsep HPD (*Highest Posterior Density*) dapat diperoleh estimator Bayesian yang merupakan mean atau modus dari distribusi posterior. Pada pendekatan Bayesian, dengan data yang observasi x dan parameter model untuk data adalah θ . Teorema Bayes digunakan untuk menentukan distribusi θ dengan syarat x , yaitu:

$$p(\theta | x) = \frac{l(x | \theta)p(\theta)}{p(x)}, \quad (2.11)$$

Parameter prior θ atau $p(\theta)$ diupdate dengan menggunakan informasi sampel yang terdapat dalam *likelihood* data atau $l(x|\theta)$, untuk memperoleh informasi *posterior* $p(\theta|x)$ yang akan digunakan dalam pengambilan keputusan. Dengan $p(x)$ adalah *normalized constant*, *posterior* dalam persamaan (2.12) dapat dituliskan sebagai berikut:

$$p(\theta | x) \propto p(\theta)l(x | \theta) \quad (2.12)$$

Metode Bayesian ini diadopsi dari nama penemunya, yaitu Thomas Bayes (1702-1761). Metode ini mulai dikenal pada tahun 1764 setelah Thomas Bayes meninggal Dunia. Meskipun demikian, metode ini mulai sering digunakan seiring dengan perkembangan teknologi dan penggunaan komputer pada abad 20 (Ismartini, 2013).

2.8.1 Distribusi Prior

Dalam pendekatan Bayesian, selain mempertimbangkan fungsi *likelihood*, penentuan distribusi prior sangat penting sehingga diperoleh distribusi posterior yang proper. Secara umum jenis prior yang sering digunakan dalam pendekatan Bayesian adalah:

1. *Conjugate prior* atau *non conjugate prior* (Ntzoufras, 2009), yaitu prior yang berkaitan atau bersesuaian dengan fungsi *likelihood*.
2. *Informative prior* atau *Non-Informative Prior* (Ntzoufras, 2009), merupakan prior yang terkait dengan diketahui atau belum diketahuinya pola/ frekuensi distribusi dari data.

3. *Proper prior* atau *Improper prior* (*Jeffreys prior*) (Ntzoufras, 2009), merupakan prior yang terkait dengan pemberian bobot/densitas di setiap titik apakah terdistribusi secara uniform atau tidak.
4. *Pseudo Prior* (Carlin dan Chib, 1995) yaitu prior yang berkaitan dengan hasil elaborasi dengan pendekatan statistika klasik.
5. *Reference Prior* diusulkan oleh Jose Bernardo di dalam papernya pada tahun 1979, kemudian dikembangkan lebih lanjut oleh Jim Berger dan lainnya dari tahun 1980 sampai dengan sekarang. Mereka membawa kembali *objective Bayesian*. *Objective Bayesian* memodelkan data secara probabilistik. Merepresentasikan sesuatu yang tidak pasti dari hal yang tidak diketahui dari data menggunakan peluang distribusi *noninformative prior* (Berger, 2005).

Ide dari *reference prior* ini adalah untuk memformalkan *noninformative prior*. Yaitu sebuah fungsi yang memaksimalkan beberapa ukuran jarak atau mendivergenkan antara posterior dan prior. Beberapa ukuran divergen dapat digunakan, seperti Kullback-Leibler divergen atau jarak Hellinger. Dengan memaksimalkan divergen, kita memungkinkan data untuk dapat memaksimumkan efek estimasi posterior. Untuk parameter satu dimensi, dapat disimpulkan *reference prior* dan *Jeffrey prior* adalah ekuivalen. Bila multidimensi maka keduanya berbeda (Jordan, 2010).

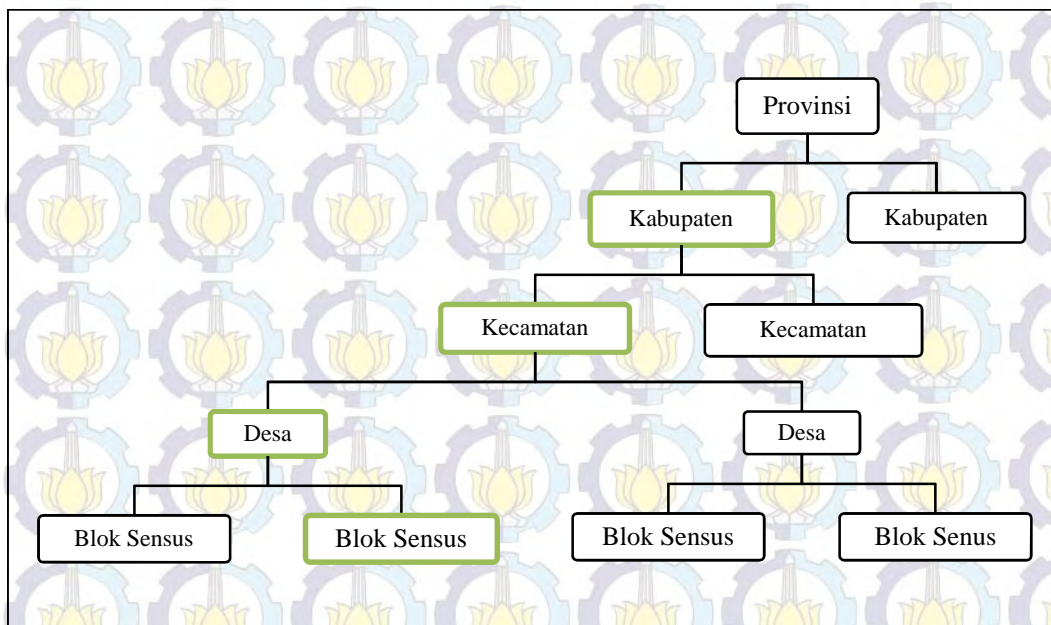
Penentuan jenis prior yang akan digunakan selain subyektifitas peneliti, juga mempertimbangkan ketersediaan informasi dari penelitian-penelitian sebelumnya yang sejenis.

2.8.2 Hierarchical Bayes dalam Reference Prior Sampling

Data hirarki merupakan data yang memiliki tingkatan dari yang sederhana hingga yang kompleks. Metode Bayesian telah mengakomodasi analisis dari data hirarki. Model hirarki Bayes adalah model hirarki dengan menggunakan metode Bayesian dalam proses estimasi parameter (Allenby dkk., 2005). Pada estimasi parameter dengan model hirarki Bayes, proses hirarki terdapat pada saat penentuan prior yang memiliki dua tahap, yaitu penentuan prior pada model

makro dan penentuan prior lainnya (Raudenbusch dan Bryk, 2002 dalam Ismartini, 2013).

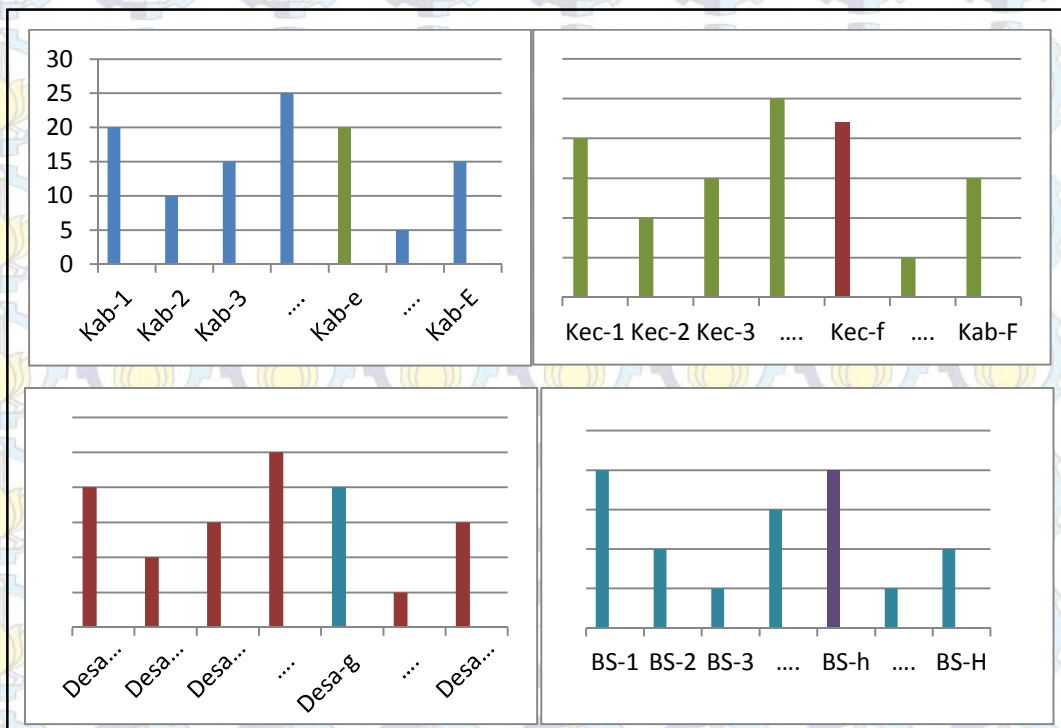
Penarikan sampel dengan *reference prior sampling* hanya menggunakan prinsip struktur hirarki tersebut pada data yang telah ada, sebagaimana Gambar 2.2. Proses penarikan sampel dimulai dengan pemilihan pada hirarki tertinggi pada provinsi, yaitu kabupaten, hingga hirarki terendah, yaitu blok sensus. Pada proses penarikan sampel ini semua kabupaten akan terpilih pada level pertama. Dari setiap kabupaten, beberapa kecamatan akan terpilih pada level kedua. Dari setiap kecamatan terpilih, akan terpilih satu atau beberapa desa pada level ketiga. Dari desa terpilih, akan terpilih satu atau beberapa blok sensus. Artinya bahwa suatu blok sensus tidak akan terpilih jika hirarki di atasnya tidak terpilih.



Gambar 2.2 Skema Hirarki dalam *Reference Prior Sampling*

Pemilihan pada setiap hirarki didasarkan pada distribusi *reference prior* pada hirarki tersebut. Distribusi *reference prior* tersebut didasarkan pada data yang telah ada. Dalam konsep *objective Bayesian*, yaitu memodelkan data secara probabilistik (Berger, 2005). Distribusi data menunjukkan peluang dari beberapa wilayah pada suatu hirarki. Distribusi pada hirarki ini akan memiliki hubungan dengan distribusi *reference prior* pada hirarki di bawahnya. Ilustrasi penarikan

sampel dengan *reference prior* ini dapat dilihat pada Gambar 2.3. Misalkan dari distribusi pada hirarki kabupaten, terpilih kabupaten Kab-e. Pemilihan Kab-e ini dilakukan dengan membuat distribusi kumulatif dari data variabel yang dijadikan *reference prior*, kemudian dibangkitkan suatu bilangan random berdistribusi Uniform (1, total kumulatif). Jika bilangan random \leq kumulatif kabupaten Kab-e maka dilanjutkan ke kabupaten selanjutnya, jika tidak maka kabupaten Kab-e tersebut terpilih sebagai sampel. Pada kabupaten Kab-e, terdapat distribusi pada hirarki kecamatan dan kemudian dipilih kecamatan Kec-f dengan cara yang sama sebagaimana pemilihan sampel kabupaten. Pada kecamatan Kec-f, terdapat distribusi pada hirarki desa dan kemudian terpilih desa Desa-g. Pada desa Desa-g, terdapat distribusi pada hirarki blok sensus, kemudian terpilih blok sensus BS-h.



Gambar 2.3 Skema Penarikan Sampel dengan Struktur Hirarki dalam *Reference Prior Sampling*

2.8.3 Markov Chain Monte Carlo

Metode *Markov Chain Monte Carlo* (MCMC) pertama kali digunakan oleh Metropolis untuk membuktikan teori fisika pada tahun 1950. Seiring dengan perkembangan teknologi komputer, metode ini mulai populer. MCMC merupakan

metode yang sering digunakan dalam pemodelan dengan pendekatan Bayesian karena distribusi prior yang ada memerlukan proses integrasi dengan dimensi yang besar sehingga proses penghitungannya menjadi sangat rumit (Carlin dan Chib, 1995).

MCMC bekerja dengan membangkitkan data parameter sesuai proses Markov Chain dengan menggunakan simulasi Monte Carlo secara iteratif hingga memperoleh distribusi posterior yang stasioner. Hal ini berbeda dengan teknik simulasi langsung yang tidak dapat digunakan untuk membangkitkan sampel dari berbagai bentuk distribusi *posterior* yang ada. Selain itu, MCMC lebih bersifat umum dan fleksibel (Ntzoufras, 2009).

Penggunaan MCMC dapat mempermudah analisis sehingga keputusan yang diambil dari hasil analisis akan dapat dilakukan dengan cepat dan tepat. Ada dua kemudahan yang diperoleh dari penggunaan metode MCMC pada analisis Bayesian (Iriawan, 2003). Pertama, metode MCMC dapat menyederhanakan bentuk integral yang kompleks dengan dimensi besar menjadi bentuk integral yang sederhana dengan satu dimensi. Kedua, dengan menggunakan metode MCMC, estimasi densitas data dapat diketahui dengan cara membangkitkan data mengikuti pola suatu rantai markov yang berurutan sebanyak N .

Ismartini (2013) menyebutkan bahwa salah satu hal penting dari proses MCMC adalah melihat apakah proses estimasi telah mencapai kondisi ekuilibrium. Hal ini dilakukan untuk menjamin bahwa sampel parameter yang diperoleh merupakan sampel dari distribusi posterior target parameter. Kondisi ekuilibrium tercapai bila sampel telah memenuhi sifat dari Markov Chain yang *strongly ergodic* (Boldstat, 2010). Sifat tersebut di antaranya:

- a. *Irreducible*, artinya selama iterasi estimasi parameter proses MCMC memiliki pergerakan nilai yang acak atau sampel parameter yang dibangkitkan melalui proses MCMC bersifat random.
- b. *Aperiodic*, artinya selama iterasi estimasi parameter proses MCMC tidak mempunyai periode tertentu untuk pasti kembali ke kondisi semula atau sampel parameter yang dibangkitkan tidak memiliki pola periodik dalam domain nilai tertentu.

- c. *Recurrent*, artinya perubahan sampel parameter terjadi secara stabil dalam domain nilai tertentu.

2.8.4 Gibbs Sampler

Metode *Gibbs Sampler* dikembangkan pertama kali dalam penelitian di bidang komputer grafik pada tahun 1984 di Amerika Serikat oleh S. Geman dan D. Geman. *Gibbs Sampler* merupakan generator yang sangat efisien, sehingga sering digunakan sebagai generator variabel-variabel random pada analisis data yang menggunakan metode MCMC. Metode *Gibbs Sampler* dapat didefinisikan sebagai suatu teknik simulasi untuk membangkitkan variabel *random* dari suatu distribusi tertentu secara tidak langsung menggunakan prinsip distribusi bersyarat penuh, tanpa harus menghitung fungsi densitas dari suatu distribusi data (Casella dan George, 1992). Dalam banyak kasus, terlihat bahwa cara kerja *Gibbs Sampler* hanya didasarkan pada sifat dasar rantai *markov*. Secara detail metode ini dapat diterangkan sebagai berikut (Geman dan Geman, 1984) dan (Iriawan, 2003):

Bila terdapat sebanyak $b+1$ variabel random yaitu θ, x_1, \dots, x_b , tergabung dalam suatu distribusi gabungan,

$$f(\theta, x_1, \dots, x_b) \quad (2.13)$$

dimana distribusi bersyarat:

$$\begin{aligned} f(\theta | x_1, \dots, x_b) \\ f(x_1 | \theta, x_2, \dots, x_b) \\ f(x_b | \theta, x_1, \dots, x_{b-1}) \end{aligned} \quad (2.14)$$

dianggap telah diketahui sebagai bentuk *full-conditional* dari persamaan (2.13).

Untuk mendapatkan karakteristik distribusi dari θ atau sering dikatakan sebagai distribusi *marginal* $f(\theta)$, maka harus dilakukan proses integral rangkap sebanyak variabel *random* yang tersisa dalam distribusi gabungannya, yaitu:

$$f(\theta) = \int \dots \int f(\theta, x_1, \dots, x_b) d\theta dx_1 \dots dx_b \quad (2.15)$$

Pengintegralan dalam persamaan (2.15) jelas akan sangat sulit atau bahkan tidak mungkin dilakukan jika fungsi gabungannya sangat kompleks. Untuk

menyelesaikan permasalahan ini, dapat menggunakan metode *Gibbs Sampler* dengan syarat distribusi bersyarat penuh (*full conditional*) setiap variabel yang terkandung di dalamnya telah diketahui. Dengan metode ini, tanpa harus menghitung dan mengetahui bagaimana bentuk fungsi *marginal*-nya, karakteristik setiap *marginal* variabel *random* yang berada di dalam fungsi gabungan akan dapat diketahui.

Pendekatan analisis karakteristik distribusi *marginal* dengan *Gibbs Sampler* ini dilakukan melalui pembangkitan data secara numerik dengan simulasi *Monte Carlo* pada distribusi bersyarat variabel *random* yang sedang dipelajari terhadap semua sisa variabel *random* yang ada di dalam fungsi distribusi gabungannya (Iriawan, 2003).

Pada persamaan (2.13) ingin diketahui karakteristik distribusi *marginal* $f(\theta)$, maka *Gibbs Sampler* dengan metode *Monte Carlo* akan membangkitkan beberapa data θ yang berdistribusi $f(\theta | x_1, \dots, x_b)$ dan karakteristik distribusi *marginal*-nya. Persamaan (2.15) diestimasi berdasarkan pada data-data hasil simulasi *Monte Carlo* tersebut. Hal ini menjelaskan bahwa dengan data yang sangat besar jumlahnya, maka nilai-nilai yang diperoleh berdasarkan data tersebut akan mencerminkan kondisi suatu populasinya.

Pada penelitian ini, variabel X tidak dibangkitkan pada proses *Gibbs Sampler* karena menggunakan data dari sampel yang telah diperoleh. Dengan demikian, tahapan proses pembangkitan data setiap variabel *random* dalam fungsi kepadatan gabungan dapat dituliskan ke dalam bentuk Algoritma 2.3 sebagai berikut:

Algoritma 2.3

1. Tentukan nilai awal $\theta_0^{(0)}, \theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_b^{(0)}, X$
2. Bangkitkan,
 - Tahap 1. $\theta_0^{(1)}$ dari $f(\theta_0 | \theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_b^{(0)}, X)$
 - Tahap 2. $\theta_1^{(1)}$ dari $f(\theta_1 | \theta_0^{(1)}, \theta_2^{(0)}, \dots, \theta_b^{(0)}, X)$

Tahap 3. $\theta_2^{(1)}$ dari $f(\theta_2 | \theta_0^{(1)}, \theta_1^{(1)}, \dots, \theta_b^{(1)}, X)$

\vdots

Tahap $b+1$. $\theta_b^{(1)}$ dari $f(\theta_b | \theta_0^{(1)}, \theta_1^{(1)}, \dots, \theta_{b-1}^{(1)}, X)$

Sehingga diperoleh nilai data bangkitan pada iterasi pertama yaitu $\theta_0^{(1)}, \theta_1^{(1)}, \theta_2^{(1)}, \dots, \theta_b^{(1)}$.

3. Menggunakan hasil bangkitan pada iterasi pertama sebagai nilai awal pada iterasi kedua. Kemudian bangkitkan data sesuai tahapan berikut ini:

Tahap 1. $\theta_0^{(2)}$ dari $f(\theta_0 | \theta_1^{(1)}, \theta_2^{(1)}, \dots, \theta_b^{(1)}, X)$

Tahap 2. $\theta_1^{(2)}$ dari $f(\theta_1 | \theta_0^{(2)}, \theta_2^{(1)}, \dots, \theta_b^{(1)}, X)$

Tahap 3. $\theta_2^{(2)}$ dari $f(\theta_2 | \theta_0^{(2)}, \theta_1^{(2)}, \dots, \theta_b^{(1)}, X)$

\vdots

Tahap $b+1$. $\theta_b^{(2)}$ dari $f(\theta_b | \theta_0^{(2)}, \theta_1^{(2)}, \dots, \theta_{b-1}^{(2)}, X)$

Sehingga diperoleh nilai data bangkitan pada iterasi kedua yaitu $\theta_0^{(2)}, \theta_1^{(2)}, \theta_2^{(2)}, \dots, \theta_b^{(2)}$.

Lakukan proses tersebut hingga iterasi ke- N atau sampai konvergen.

Menggunakan hasil bangkitan pada iterasi ke $N - 1$ sebagai nilai awal pada iterasi ke- N . Kemudian membangkitkan tahap berikut ini:

Tahap 1. $\theta_0^{(N)}$ dari $f(\theta_0 | \theta_1^{(N-1)}, \theta_2^{(N-1)}, \dots, \theta_b^{(N-1)}, X)$

Tahap 2. $\theta_1^{(N)}$ dari $f(\theta_1 | \theta_0^{(N)}, \theta_2^{(N-1)}, \dots, \theta_b^{(N-1)}, X)$

Tahap 3. $\theta_2^{(N)}$ dari $f(\theta_2 | \theta_0^{(N)}, \theta_1^{(N)}, \dots, \theta_b^{(N-1)}, X)$

\vdots

Tahap $b+1$. $\theta_b^{(N)}$ dari $f(\theta_b | \theta_0^{(N)}, \theta_1^{(N)}, \dots, \theta_{b-1}^{(N)}, X)$

Sehingga diperoleh nilai data bangkitan pada iterasi ke- N yaitu $\theta_0^{(N)}, \theta_1^{(N)}, \theta_2^{(N)}, \dots, \theta_b^{(N)}$.

4. Berdasarkan N buah data yang sudah dibangkitkan, kemudian diestimasi karakteristik setiap fungsi kepadatan *marginal* $f(\theta_0), f(\theta_1), \dots, f(\theta_b)$.

Metode *Gibbs Sampler* sering dinamakan sebagai Metode MCMC. Karena sifat dan karakteristik pembangkitan data randomnya sangat erat sekali dengan sifat *Markov Chain*. Sifat tersebut adalah bahwa *Gibbs Sampler* membangkitkan data random pada setiap iterasi atau bangkitan data ke- i sangat tergantung sekali dengan nilai *random* yang pernah dibangkitkan satu iterasi sebelumnya dan tidak tergantung pada nilai random pada iterasi kedua, ketiga hingga ke- N yang sebelumnya (Iriawan, 2003).

Berikut merupakan ilustrasi penerapan *Gibbs Sampler* pada penelitian ini (dengan $x=X_2$ dan $m=X_1$). Pada estimasi proporsi pengangguran, digunakan konsep variabel random x (jumlah pengangguran) berdistribusi binomial (m, θ). Dengan sampel beberapa blok sensus dengan data x_d dan m_d , fungsi kepadatan bersama mengikuti persamaan (2.16).

$$f(\theta, x_d, m_d) = \binom{m_d}{x_d} \theta^{x_d} (1-\theta)^{m_d-x_d}, \quad (2.16)$$

Keterangan:

$x_d = 0, 1, \dots, m_d$

d = sampel ke $1, \dots, n$

Bila diketahui data variabel random x dan m dari sampel sebagai likelihood ($f(x, m | \theta)$) dan digunakan prior *conjugate* dengan informasi dari data populasi ($f(\theta)$), maka selanjutnya nilai parameter θ dapat diestimasi. Estimasi dapat dilakukan dengan mengikuti alur Algoritma 2.3. Iterasi pertama dilakukan sehingga diperoleh $\theta^{(1)}$, dan pada iterasi kedua diperoleh $\theta^{(2)}$. Iterasi dilakukan hingga ke N dan diperoleh $\theta^{(N)}$. Dari N buah bilangan random yang sudah dibangkitkan, kemudian diestimasi karakteristik fungsi kepadatan *marginal* $f(\theta)$.

2.8.5 Peran Bayesian dalam Survey Sampling

Brewer (2013) menjelaskan bahwa survei sampling mulai diperkenalkan oleh Kiaer pada tahun 1895, 1899, dan 1901. Namun, peneliti pertama yang memberikan pembuktian secara teoritis adalah Bowley pada tahun 1906. Bowley mengusulkan teori inferensi pada survei sampling dengan menggunakan *Central*

Limit Theorem versi Bayesian milik Edgeworth. Dia mampu menilai akurasi dari estimator untuk sampel besar yang didapatkan dengan sampel acak sederhana dari populasi yang besar. Pada tahun 1912, Bowley berhasil mendemonstrasikan sebuah survei. Pada tahun 1926, Bowley dalam monografinya menyampaikan pendekatan *purposive* dan *random* dalam pemilihan sampel. Dia menyarankan agar dilakukan pengembangan pada kedua pendekatan tersebut.

Neyman (1934) mengeluarkan paper yang mengkhususkan pada randomisasi. Dia mengkritik penggunaan sampling *purposive* oleh Gini dan Galvani pada tahun 1929. Dengan menggunakan sampling tersebut, mereka tidak dapat menggunakan konfiden interval yang juga dirumuskan pada paper tersebut. Neyman juga menolak konsep Bayesian yang digunakan Bowley. Pada dua atau tiga dekade berikutnya, tidak terlihat peran Bayesian dalam survey sampling (Brewer, 2013).

Dalam desain sampling, Basu (1968) menjelaskan bahwa rencana penarikan sampel bukanlah materi dari analisis Bayesian. Hal ini disebabkan distribusi posterior dari parameter yang diestimasi bersifat independen dengan aturan sampling. Meskipun demikian, Zack (1969) menjelaskan bahwa set data yang berbeda kemungkinan akan membawa inferensi Bayesian pada presisi yang berbeda. Dia menggunakan distribusi prior untuk memilih sampel optimal dengan jumlah sampel yang telah ditentukan. Pemilihan optimal dengan Bayes yang dilakukannya bersifat tidak *random, without replacement*, dan dilakukan secara sekuensial. Secara umum, pemilihan sekuensial optimal ini sulit dilakukan. Praktisi dalam banyak kasus tidak dapat menentukan joint distribusi prior pada vektor dari N unit populasi (Zack, 1970). Hal ini sejalan dengan yang disampaikan Meeden (2012), bahwa pendekatan Bayesian ini belum pernah dipraktekkan. Permasalahannya adalah sulit untuk menemukan distribusi prior yang memungkinkan penggunaan informasi prior dari populasi.

Bayesian juga berkontribusi pada alokasi sampel dalam penarikan sampling berstrata dan penentuan jumlah sampel. Zack (1970) menyebutkan bahwa alokasi optimal Bayes telah dipelajari Aggarwal (1959, 1966), Ericson (1965, 1968), Zack (1967), Guttman dan Draper (1968), dan Grosh (1969). Aggarwal dan Ericson mempelajari pada stratified dengan fase tunggal,

sedangkan Zack, Guttman dan Draper serta Grosh mempelajari pada stratified dengan fase ganda. Guttman dan Draper (1968) mengusulkan alokasi sample optimal dengan menggunakan satu karakteristik yang didapatkan dari fase pertama. Beberapa bulan berikutnya mengembangkan dengan menggunakan beberapa karakteristik.

Dalam survei biasanya terdapat data yang missing (*missing data*). Di sini bayesian memberikan solusi untuk mengatasi hal tersebut. Haziza (2009) menjabarkan beberapa solusi Bayesian dalam mengatasi *missing data*, yaitu: Rubin (1987) mengusulkan penggunaan *Approximate Bayesian Bootstrap (ABB)*, Schafer (1997) serta Little dan Rubin (2002) mempelajari mengenai *Bayesian Proper Imputation*. Kott (1995) dan Kim dkk (2006) merumuskan *Bayesian Imputation Model*.

Rao (2011) menerangkan bagaimana peran Bayesian dalam inferensi dari data survey. Royal dan Pfeffermann (1982) dalam Rao (2011) mempelajari inferensi Bayesian dalam mean populasi dengan asumsi normal dan *flat (diffuse) prior* pada parameter dari model regresi linier. Scott dan Smith (1969) dalam Rao (2011) mendapatkan posterior mean dan posterior varians mean populasi di bawah model linier dengan efek random, normalitas, dan *diffuse prior* pada parameter model. Sedransk (1977) dalam Rao (2011) mempelajari model regresi dengan slope random menggunakan distribusi prior. Dengan mengikuti pendekatan Scott dan Smith (1969) dalam Rao (2011), dia menerapkan pada data bank dari U.S. Federal Reserve Board untuk mengestimasi *current monetary total*. Dihasilkan inferensi posterior yang lebih efisien dibandingkan dengan menggunakan *noninformative prior*. Malec dan Sedransk (1985) dalam Rao (2011) mengembangkan hasil dari Scott dan Smith untuk sampling tiga tahap. Nandram, dkk. (1997) dalam Rao 2011 menerapkan pendekatan Bayesian untuk mendapatkan estimator *order restricted* dari komposisi umur populasi dengan metode MCMC. Sedransk (2008) dalam Rao (2011) mendaftar kemungkinan penggunaan metode parametrik Bayesian untuk survey sample, termasuk estimasi dari survey, alokasi sampel optimal, dan *Small Area Estimation*. Pfeffermann, dkk. (2006) menyampaikan penerapan pendekatan Bayesian untuk estimasi dari multilevel model di bawah *informative sampling* di mana metode *frequentist* sulit

untuk menerapkannya. Di sisi lain, Rao menjelaskan bahwa dengan penggunaan noninformative prior pada parameter model dan menerapkan MCMC akan lebih efisien dan mudah dalam menangani model sampel yang kompleks.

Rao (2011) menjelaskan bahwa pada untuk survey dengan skala besar dan *multipurpose*, metode Bayesian parametrik dengan berbasis asumsi distribusi memiliki keterbatasan karena sulit dalam memvalidasi asumsi parametrik. Oleh karena itu, perlu digunakan pendekatan Bayesian nonparametric. Akan tetapi, diperlukan spesifikasi fungsi likelihood nonparametric pada seluruh data sampel dan sebuah distribusi prior pada vector parametric. Fungsi likelihood dengan basis seluruh data sampel merupakan noninformative karena semua kemungkinan nilai observasi dari vektor parameter memiliki fungsi likelihood yang sama (Godambe, 1969 dalam (Rao, 2011)). Salah satu cara mengatasi kesulitan ini yaitu dengan mengasumsikan sebuah prior *informative (exchangeable)* pada *N-dimensional* vektor parameter dan mengkombinasikannya dengan *noninformative likelihood* (Ericson, 1969). Hal ini untuk mendapatkan posterior informative, tetapi inferensi tidak bergantung pada desain sampel. Meeden dan Vardemen (1991) dalam Rao (2011) menggunakan Polya posterior (PP) atas data yang tidak terobservasi dengan mengasumsikan *unseen like the seen* (ekuivalen dengan *exchangeability*). Pada kasus ini posterior tidak berasal dari distribusi prior tunggal (Meeden, 1995 dalam (Rao, 2011)), ini kemudian disebut sebagai *pseudo-posterior*. Hal ini juga mirip dengan *Bayesian bootstrap* (Rubin, 1981 dalam (Rao, 2011)). Meeden (1995) dalam Rao (2011) mengembangkan pendekatan PP dengan menggunakan informasi tambahan dari populasi dengan membuat sebuah asumsi prior yang kuat dengan rasio $r_i = y_i/x_i$. Meeden (1999) dalam Rao (2011) mempelajari sampling berkelompok dua tahap dan PP berbasis dua tahapnya menghasilkan posterior mean dan posterior varians yang mendekati hasil *design-based*.

Penelitian yang paling terkini dalam inferensi survey sampling dengan pendekatan Bayesian adalah dalam *Small Area Estimation* (SAE). Datta (2009) menjabarkan bahwa SAE dengan pendekatan Bayesian ini diawali oleh Datta dan Ghosh (1991) dalam Datta (2009) dengan mengusulkan *Bayesian formulation* dalam SAE. Kemudian Stroud (1991) dalam Datta (2009) mengusulkan *Hierarchical Bayesian* (HB) untuk SAE dengan respon data biner. Stroud (1994)

dalam Datta (2009) kemudian mengembangkan pendekatan Bayesian untuk data biner ini pada berbagai desain sampling. Ghosh (1998) dalam Datta (2009) menggunakan model HB untuk SAE dalam *General Linier Model* (GLM). Rao (2011) menyebutkan bahwa HB memberikan kemampuan untuk mengatasi model yang kompleks dan memberikan estimasi yang pasti.

Raghunathan dan Rubin (1997) menyebutkan bahwa pendekatan berdasar randomisasi yang dibawa oleh Neyman (1934) sangat dominan digunakan dalam survey sampling dan cocok dalam mengevaluasi bagaimana prosedur statistik dioperasikan dalam jangka panjang. Di sisi lain, pendekatan Bayesian cocok pada tugas memberikan pedoman membuat prosedur yang sulit. Kedua pendekatan bersifat komplemen dan tidak bersaing.

2.9 Ketenagakerjaan

Data tentang situasi ketenagakerjaan merupakan salah satu data pokok yang menggambarkan kondisi perekonomian, social, dan tingkat kesejahteraan penduduk di suatu wilayah dalam kurun waktu tertentu (BPS, 2011). Untuk mendapatkan data mengenai ketenagakerjaan, BPS melakukan Survei Angkatan Kerja Nasional (Sakernas). Karakteristik utama ketenagakerjaan yang dikumpulkan di dalam Sakernas meliputi, jumlah penduduk yang bekerja, jumlah pengangguran, tingkat pengangguran terbuka (TPT), lapangan usaha pekerja, jumlah jam kerja, status dalam pekerjaan utama, dsb.

Meskipun demikian, dari survey tersebut, TPT merupakan informasi yang mungkin paling bernilai ekonomi dan politis. Informasi ini menggambarkan bagaimana kemampuan pemerintah dalam menciptakan lapangan pekerjaan bagi angkatan kerja yang setiap tahun bertambah.

2.9.1 Bekerja

Bekerja adalah kegiatan melakukan pekerjaan dengan maksud memperoleh atau membantu memperoleh penghasilan atau keuntungan paling sedikit selama satu jam dalam seminggu terakhir. Bekerja selama satu jam tersebut harus dilakukan berturut-turut dan tidak terputus. Penghasilan atau keuntungan mencakup upah/ gaji termasuk semua tunjangan dan bonus bagi pekerja/

karyawan/ pegawai dan hasil berupa sewa atau keuntungan, baik berupa uang atau barang termasuk bagi pengusaha (BPS, 2011).

2.9.2 Konsep Pengangguran dalam Sensus dan Survei BPS

Krismawati dan Agustiyani (2010) menjelaskan bahwa pengangguran dapat didefinisikan secara standar ataupun secara luas (*relaxed*). Pengangguran definisi standar yaitu meliputi penduduk yang tidak bekerja tetapi sedang mencari pekerjaan/ mempersiapkan suatu usaha. Sedangkan pengangguran definisi luas juga mencakup penduduk yang tidak aktif mencari kerja tetapi bersedia/ siap bekerja.

Sejak tahun 2001, definisi pengangguran yang digunakan oleh Sakernas adalah definisi luas, sehingga pengangguran mencakup empat kriteria yaitu: mencari pekerjaan, mempersiapkan usaha, putus asa/merasa tidak mungkin mendapatkan pekerjaan (*discouraged worker*) dan sudah diterima bekerja tapi belum mulai bekerja.

Pengangguran di SP 2010 ditanyakan secara ringkas dibandingkan dengan Sakernas. Pertanyaan "apakah mencari kerja?" dan "apakah mempersiapkan usaha?" menjadi satu pertanyaan "apakah mencari kerja/mempersiapkan usaha?" Pertanyaan mengenai penduduk yang tidak mencari kerja karena merasa putus asa dan yang sudah diterima bekerja tetapi belum mulai bekerja diringkas menjadi satu pertanyaan "apakah bersedia bekerja apabila ada yang menyediakan?" Oleh karena itu dari SP 2010 tidak dapat mengukur angka setengah pengangguran, pekerja berdasarkan jenis pekerjaan, mobilitas pekerja seperti Sakernas, tingkat upah, dan lain-lain. Oleh karena perbedaan cakupan, bentuk pertanyaan dan kualitas petugas maka data ketenagakerjaan SP2010 juga tidak dapat dibandingkan secara langsung dengan Sakernas. Hasil sensus penduduk hanya dapat dibandingkan dengan sensus penduduk terdahulu.

Indikator yang biasa digunakan untuk mengukur pengangguran adalah Tingkat Pengangguran Terbuka (TPT). TPT memberikan indikasi besarnya penduduk usia kerja yang termasuk dalam pengangguran. TPT merupakan persentase pengangguran terhadap angkatan kerja, sesuai dengan persamaan (2.17).

$$TPT = \frac{\text{Jumlah Pengangguran}}{\text{Jumlah Angkatan kerja}} \times 100\% \quad (2.17)$$

Dalam menginformasikan tingkat pengangguran dari SP 2010 perlu berhati-hati, karena walaupun konsep SP2010 mengikuti Sakernas tetapi pertanyaan SP2010 diringkas, sehingga mungkin menghasilkan data yang berbeda. Akan kurang konsisten jika tingkat pengangguran pada SP2010 disandingkan dengan tingkat pengangguran Sakernas Februari 2010.

2.9.3 Distribusi dari Data Ketenagakerjaan

Pada subbab sebelumnya telah diberikan hubungan dari TPT, jumlah pengangguran, dan jumlah angkatan kerja. TPT sama dengan jumlah pengangguran dibagi jumlah angkatan kerja dibagi seratus. Bila difokuskan pada variabel jumlah pengangguran di suatu wilayah, maka variabel ini memiliki distribusi binomial/ $B(m, \theta)$ dengan parameter m adalah jumlah angkatan kerja dan parameter θ adalah proporsi jumlah pengangguran terhadap jumlah angkatan kerja.

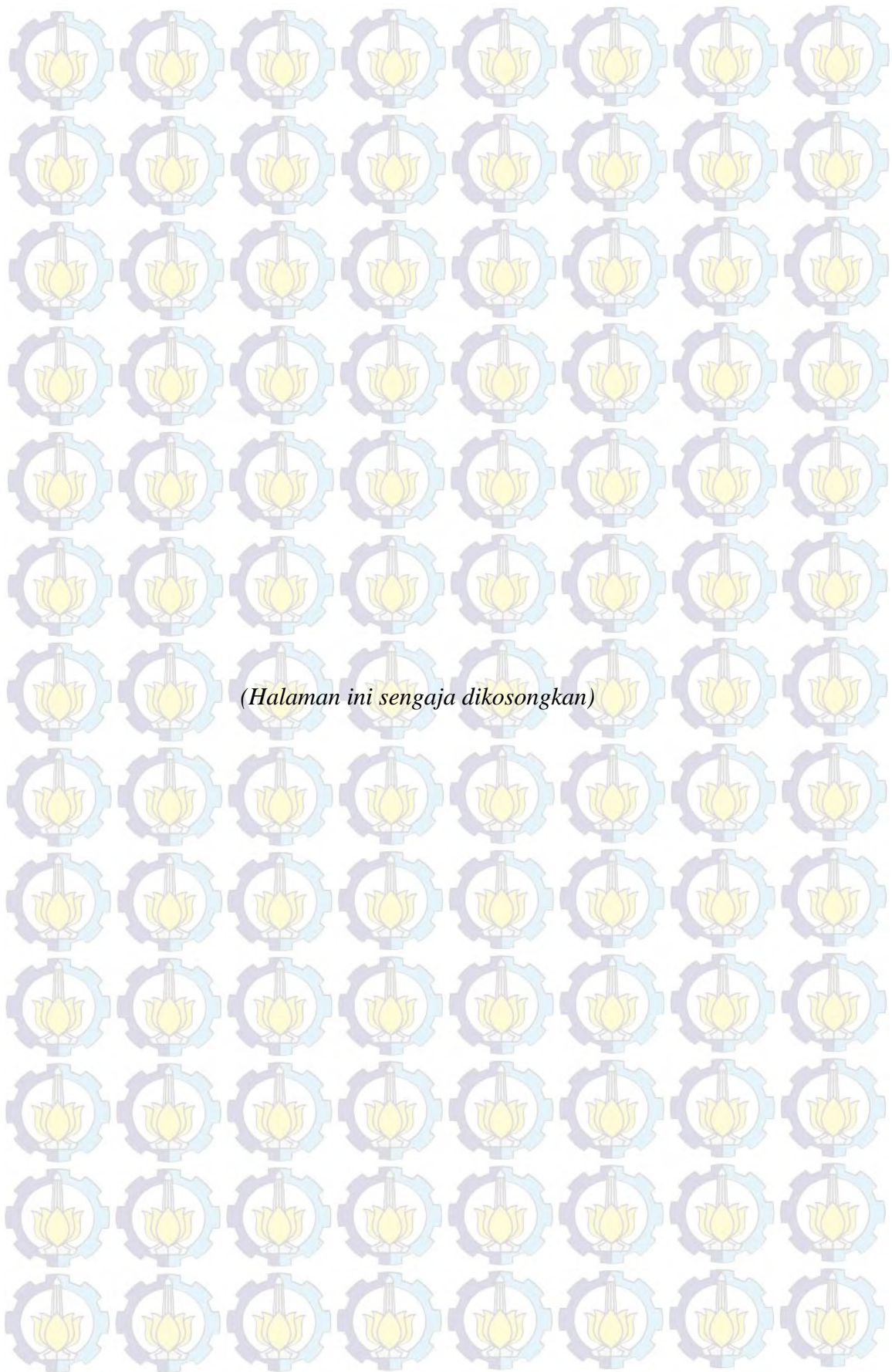
Pembuktian variabel pengangguran memiliki distribusi binomial adalah sebagai berikut:

- a. Pada individu yang tergolong angkatan kerja maka ada dua kemungkinan, yaitu bekerja atau menganggur. Dengan fokus variabel pengangguran, dapat dimisalkan individu yang menganggur dikatakan kejadian sukses (1) dan individu yang tidak menganggur dikatakan kejadian gagal (0). Pada percobaan Bernoulli, tiap percobaan memiliki dua kemungkinan yang dapat disebut kejadian sukses dan gagal serta peluang sukses θ (Congdon, 2010). Dengan demikian, kejadian tiap individu pada angkatan kerja dapat dikatakan mengikuti distribusi Bernoulli dengan parameter θ .
- b. Distribusi binomial muncul bila dihitung kejadian sukses pada m percobaan independen dan tiap kejadian menghasilkan sukses atau gagal. Bila peluang sukses sebesar θ dan total kejadian sukses adalah X , maka X mengikuti distribusi binomial (m, θ) (Congdon, 2010). Dengan kerangka konsep yang sama, bila terdapat sebanyak m individu angkatan kerja dikumpulkan di dalam satu wilayah, misalnya provinsi, kabupaten,

kecamatan, desa, ataupun blok sensus, maka dapat dianalogikan bahwa ada sebanyak kejadian m . Dengan kejadian sebanyak m dan tiap kejadian berdistribusi Bernoulli dengan parameter θ , maka pengangguran di suatu wilayah memiliki distribusi binomial dengan parameter m dan θ ($B(m, \theta)$).

Setelah diketahui bahwa jumlah pengangguran memiliki distribusi binomial, maka parameter θ ini dapat diestimasi dengan pendekatan Bayesian, yaitu MCMC dan *Gibbs Sampler*. Distribusi prior dari θ mengikuti distribusi beta-binomial karena merupakan prior *conjugate* dari binomial.

Di dalam data ketenagakerjaan, orang-orang yang bekerja dapat dikelompokkan menurut lapangan usahanya. Dengan demikian, variabel pekerja menurut lapangan usaha mengikuti distribusi multinomial. Distribusi dirichlet ini digunakan sebagai prior *conjugate* di dalam mendapatkan estimasi proporsi menurut lapangan usaha.



BAB III METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang akan digunakan menggunakan bersumber dari BPS. Data tersebut merupakan *raw data* hasil Sensus Penduduk 2010. Informasi yang digunakan seperti kode kabupaten, kode kecamatan, kode desa, kode blok sensus, jumlah angkatan kerja, jumlah pengangguran, jumlah angkatan kerja yang bekerja, dan jumlah pekerja menurut lapangan usaha.

3.2 Variabel Penelitian

Variabel utama yang mendasari penelitian ini adalah Tingkat Pengangguran Terbuka (TPT), tetapi variabel ini akan didekati dengan proksi proporsi pengangguran terhadap angkatan kerja. Selain itu, penelitian juga akan meneliti proporsi pekerja menurut lapangan usaha. Variabel yang digunakan adalah sebagai berikut:

- X_1 = jumlah angkatan kerja
- X_2 = jumlah pengangguran
- X_3 = jumlah angkatan kerja yang bekerja
- X_4 = jumlah pekerja di sektor pertanian
- X_5 = jumlah pekerja di sektor industri
- X_6 = jumlah pekerja di sektor perdagangan, rumah makan, dan jasa akomodasi.
- X_7 = jumlah pekerja di sektor jasa kemasyarakatan dan lainnya

Variabel jumlah pengangguran akan dijadikan dasar dalam penarikan sampel dengan pendekatan Bayesian dengan *Reference Prior*. Adapun untuk *Balanced Sampling* dengan *cube method* seluruh variabel akan digunakan di dalam proses penarikan sampel, ditambah data jumlah penduduk.

3.3 Definisi Operasional

Beberapa definisi yang digunakan di dalam penelitian ini adalah sebagai berikut:

- a) Penduduk Usia Kerja adalah penduduk yang berusia 15 tahun ke atas, sesuai dengan ketentuan dalam UU Ketenagakerjaan No. 13 Tahun 2003.
- b) Angkatan Kerja adalah penduduk usia kerja (15 tahun dan lebih) yang bekerja dan pengangguran.
- c) Bukan Angkatan Kerja adalah penduduk usia kerja yang pada periode referensi tidak mempunyai/melakukan aktivitas ekonomi, baik karena sekolah, mengurus rumah tangga atau lainnya (pensiun, penerima transfer/kiriman, penerima deposito/bunga bank, jompo atau alasan yang lain).
- d) Bekerja yaitu kegiatan melakukan pekerjaan dengan maksud memperoleh atau membantu memperoleh penghasilan atau keuntungan paling sedikit selama satu jam dalam seminggu yang lalu.
- e) Pengangguran adalah penduduk usia kerja yang tidak sedang bekerja atau sementara tidak bekerja, tetapi sedang mencari kerja atau mempersiapkan usaha atau bersedia bekerja apabila ada yang menyediakan.
- f) Wilayah pencacahan (wilcah) adalah bagian dari suatu wilayah desa/kelurahan yang pada umumnya merupakan wilayah kerja seorang pencacah. Wilcah harus mempunyai batas yang jelas, baik alam maupun buatan, dan diperkirakan tidak akan berubah dalam jangka waktu kurang lebih 10 tahun. Sebuah wilcah pada umumnya mencakup sekitar 200-300 rumah tangga atau bangunan fisik bukan tempat tinggal, atau gabungan dari rumah tangga dan bangunan fisik bukan tempat tinggal.
- g) Blok sensus adalah wilayah kerja pencacahan yang merupakan bagian dari suatu wilayah desa/ kelurahan. Blok sensus pada umumnya mencakup sekitar 80-120 rumah tangga.
- h) Lapanga Pekerjaan/ Usaha Utama adalah bidang kegiatan suatu organisasi/ lembaga/ usaha (*establishment*) tempat seseorang bekerja selama periode waktu acuan yang dibuat untuk data karakteristik ekonomi (atau yang dikerjakan terakhir, jika orang tersebut tidak bekerja). Dari hasil sensus penduduk 2010 terdapat 19 sektor. Pada penelitian ini 19 sektor tersebut

dibagi ke dalam empat kelompok. Sektor 1 s.d 6 dikelompokkan menjadi sektor pertanian, sektor 7-10 dikelompokkan menjadi sektor industri, sektor 11 s.d 13 dikelompokkan menjadi sektor perdagangan, rumah makan, dan akomodasi serta sektor 14 s.d 19 dikelompokkan menjadi sektor jasa kemasyarakatan dan lainnya.

3.4 Struktur Data

Sebelum masuk ke tahap penarikan sampel, data yang ada dipersiapkan sehingga memiliki struktur seperti dalam Tabel 3.1 dan Tabel 3.2.

Tabel 3.1 Struktur Data Penarikan Sampel dengan Pendekatan Bayesian (*Reference Prior*)

Pemilihan kabupaten

kode_kab	X_2

Pemilihan kecamatan

kode_kab	kode_kec	X_2

Pemilihan desa

kode_kec	kode_desa	X_2

Pemilihan blok sensus

kode_desa	kode_bs	X_2

Tabel 3.2 Struktur Data Penarikan Sampel dengan *Balanced Sampling* dengan *Cube Method*

kode_bs	jml_pddk	X_1	X_2	...	X_7

3.5 Metode Penelitian

Dalam penelitian ini metode yang digunakan adalah *Reference Prior Sampling* dan *Balanced Sampling* dengan *cube method*. Adapun untuk pengolahan menggunakan software utama R.

Secara umum rancangan sampel dalam penelitian ini adalah sebagai berikut:

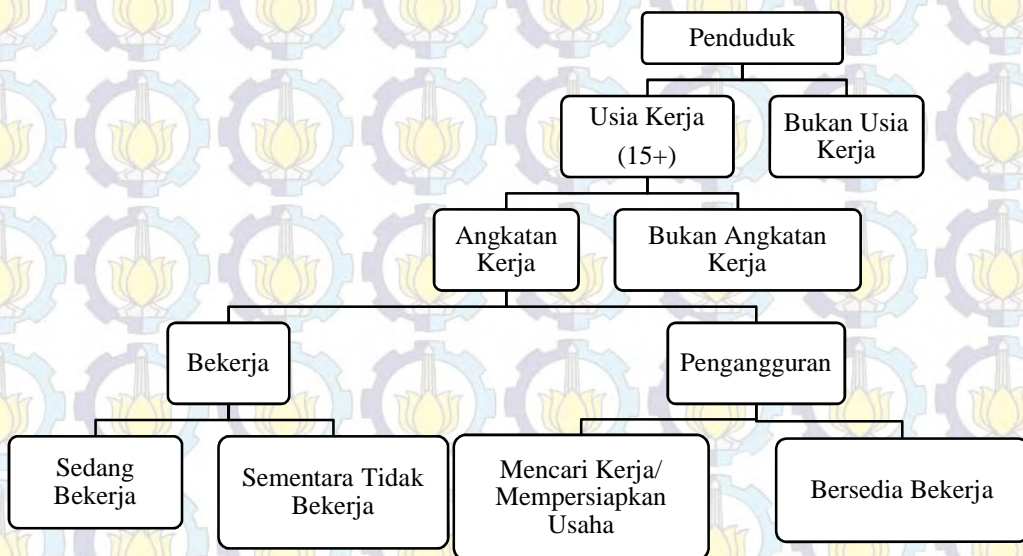
- Populasi: seluruh blok sensus di Kalimantan Tengah sesuai SP2010.
- Target populasi: seluruh blok sensus biasa di Kalimantan Tengah sesuai SP2010.
- Unit sampel, yaitu unit yang dijadikan dasar penarikan sampel: blok sensus.
- Jumlah blok sensus yang akan digunakan sebanyak 116 blok sensus. Hal ini mengikuti jumlah blok sensus yang digunakan pada Sakernas Triwulan BPS agar didapatkan perbandingan yang sebanding.

Langkah-langkah dalam penelitian ini dapat dirinci sebagaimana berikut:

- Persiapan
 - Menyiapkan data hasil pencacahan dokumen SP2010-C1 .
 - Memilih data untuk usia 15 tahun ke atas dari *raw* data sensus penduduk.
 - Mengategorikan penduduk yang bekerja didasarkan pada jawaban kuesioner pada Lampiran 1.
 - Mengategorikan penduduk yang menganggur didasarkan pada jawaban kuesioner pada Lampiran 1.
 - Mengategorikan penduduk angkatan kerja yaitu penduduk yang bekerja atau yang menganggur.
 - Mengategorikan penduduk yang bekerja menurut lapangan usahanya.

- g. Mengagregasikan keseluruhan variabel di atas menurut blok sensus.
- h. Mendapatkan kerangka sampel berupa blok sensus dengan disertai variabel tambahan seperti: jumlah pengangguran, jumlah bekerja, jumlah angkatan kerja, dan jumlah pekerja menurut lapangan usaha untuk tiap blok sensus.

Tahapan di atas mengikuti alur diagram ketenagakerjaan sesuai dengan Gambar 3.1 (Krismawati dan Agustiyani, 2010).



Gambar 3.1 Diagram Ketenagakerjaan Sensus Penduduk 2010

2. Penarikan Sampel dengan *Balanced Sampling* (Tille, 2006)

- a. *Flight Phase* dilakukan sesuai Algoritma 2.1.

Step 1 Inisialisasi

- a) Menyiapkan vektor peluang terpilih π . Unit dengan peluang terpilih mendekati 0 atau 1 dikeluarkan dari populasi sehingga yang masih ada memiliki peluang $0 < \pi_k < 1$.
- b) Menyiapkan matriks penimbang \mathbf{A} yang berisi variabel *balancing* dibagi peluang terpilih..
- c) Vektor \mathbf{r} terdiri dari $p + 1$ pertama dari π .
- d) Membuat vektor ranking $\mathbf{r} = (1 \ 2 \ \dots \ p \ p+1)'$.

e) Membuat matriks **B** terdiri dari $p + 1$ pertama kolom **A**

f) Inisial $k = p + 2$

Step 2. Proses Sampling

a) Mengambil sebuah vektor **u** dari matriks kernel **B**.

b) Hitung λ_1^* dan λ_2^* , nilai terbesar dari λ_1 dan λ_2 sehingga

$$0 \leq \psi + \lambda_1 \mathbf{u} \leq 1 \text{ dan } 0 \leq \psi + \lambda_2 \mathbf{u} \leq 1. \text{ Catat bahwa } \lambda_1^* > 0 \text{ dan } \lambda_2^* > 0.$$

c) Memilih $\psi = \begin{cases} \psi + \lambda_1 \mathbf{u} & \text{dengan peluang } q \\ \psi + \lambda_2 \mathbf{u} & \text{dengan peluang } 1 - q \end{cases}$
di mana $q = \lambda_2^* / (\lambda_1^* + \lambda_2^*)$

d) Menghapus unit dari **B** yang berkorespondensi dengan angka integer $\psi(i)$ dan diganti dengan unit baru.

e) Kembali ke Step 2.a)

Step 3. Akhir dari bagian *flight phase*

For $i = 1$ to $p+1$, DO $\pi(\mathbf{r}(i)) = \psi(i)$ ENDFOR

b. *Landing Phase* dengan *Enumerative Algorithm* dilakukan sesuai Algoritma 2.2.

a) Selesaikan $\min_{p(s)} \sum_{s \in S} p^*(s) C(s)$ pada $\sum_{s \in S} p^*(s) = 1$, $\sum_{s \in S} s p^*(s) = \pi^*$, dan

$$0 \leq p^*(s) \leq 1 \text{ untuk semua } s \in S$$

b) Bentuk list $\{s_1, s_2, \dots, s_v, \dots, s_V\}$ dari semua kemungkinan sampel dengan semua peluangnya

c) *Generate* variabel *random* u dengan sebuah distribusi uniform (0,1)

d) Pilih sampel s_j sehingga $\sum_{c=1}^{v-1} p^*(s_c) \leq u \leq \sum_{c=1}^v p^*(s_c)$

c. Mendapatkan sejumlah n sampel

d. *Quality Balancing*

3. Penarikan Sampel dengan *Reference Prior Sampling*

a. Pemilihan kabupaten

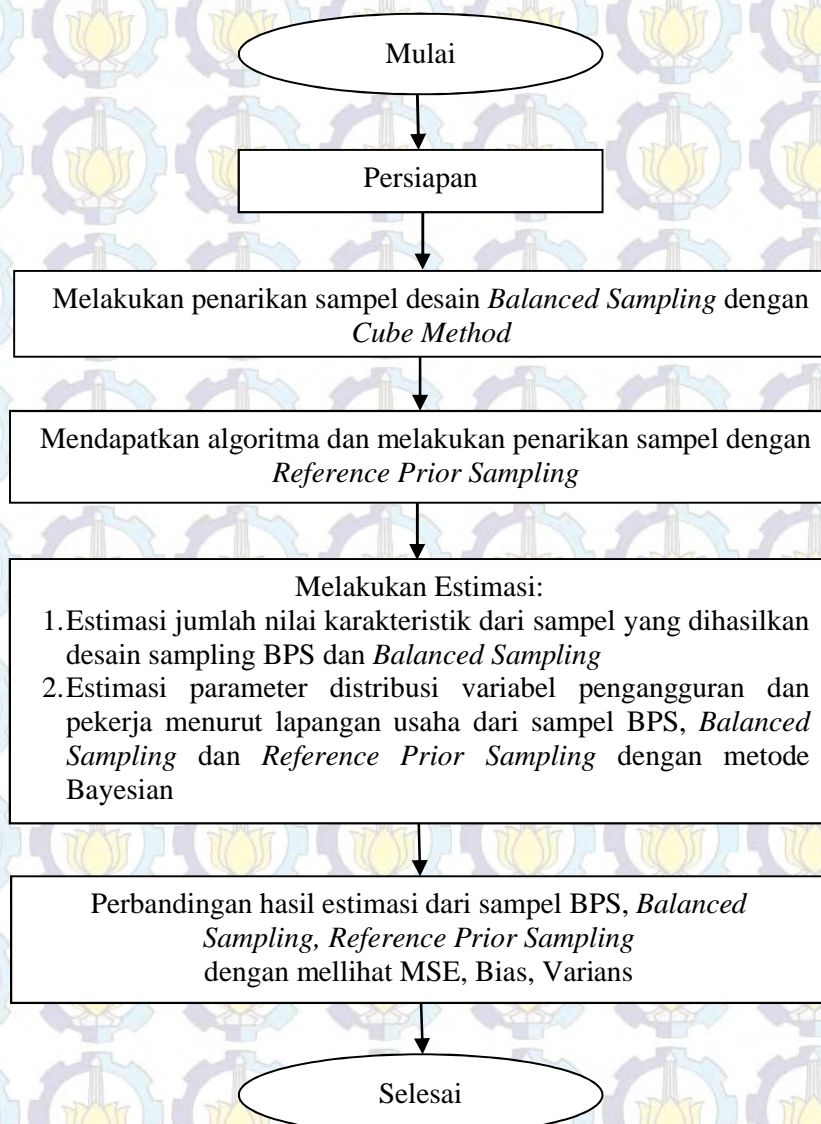
i. Membuat kumulatif jumlah pengangguran pada tiap kabupaten dalam Provinsi Kalimantan Tengah

ii. Membangkitkan U yang berdistribusi $U(1, \text{total kumulatif})$.

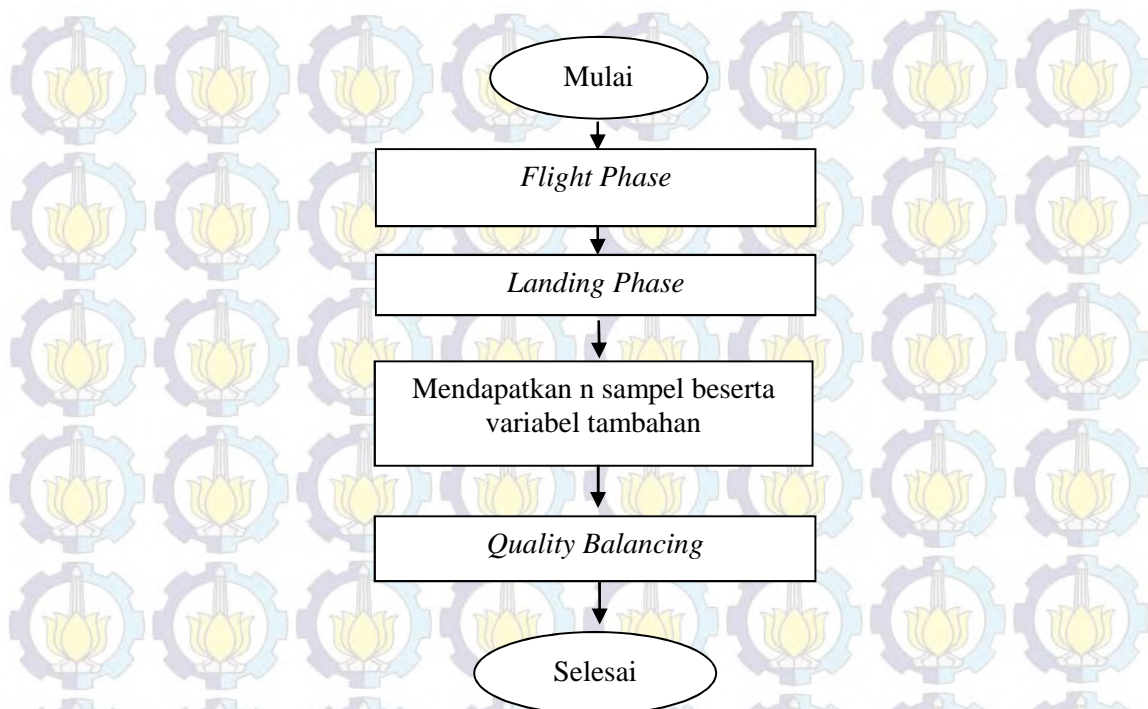
- iii. Jika $U \leq$ kumulatif kabupaten ke-e maka dilanjutkan ke kabupaten selanjutnya, jika tidak maka kabupaten tersebut terpilih sebagai sampel kabupaten ke-d.
 - b. Pemilihan kecamatan
 - i. Untuk kabupaten terpilih, dibuat kumulatif jumlah pengangguran tiap kecamatan.
 - ii. Membangkitkan V yang berdistribusi $U(1, \text{total kumulatif})$.
 - iii. Jika $V \leq$ kumulatif kecamatan ke-f maka dilanjutkan ke kecamatan selanjutnya, jika tidak maka kecamatan tersebut terpilih sebagai sampel kecamatan ke-d.
 - c. Pemilihan desa
 - i. Untuk kecamatan terpilih, dibuat kumulatif jumlah pengangguran tiap desa.
 - ii. Membangkitkan J yang berdistribusi $U(1, \text{total kumulatif})$.
 - iii. Jika $J \leq$ kumulatif desa ke-g maka dilanjutkan ke desa selanjutnya, jika tidak maka desa tersebut terpilih sebagai sampel desa ke-d.
 - d. Pemilihan blok sensus
 - i. Untuk desa terpilih, di buat kumulatif jumlah pengangguran tiap blok sensus.
 - ii. Membangkitkan K yang berdistribusi $U(1, \text{total kumulatif})$.
 - iii. Jika $K \leq$ kumulatif blok sensus ke-h maka dilanjutkan ke blok sensus selanjutnya, jika tidak maka blok sensus tersebut terpilih sebagai sampel blok sensus ke-d.
 - e. Mendapatkan sejumlah n sampel.
4. Melakukan estimasi:
- a. Total dari suatu variabel penelitian dari sampel yang dihasilkan desain sampling BPS dan desain *balanced sampling* dengan pendekatan *desain based* (Estimator Horvitz-Thompson).
 - b. Parameter distribusi dari sampel yang dihasilkan desain sampling BPS, desain *balanced sampling* dan desain pendekatan Bayesian dengan menggunakan Metode Bayesian dengan menggunakan MCMC *Gibbs Sampler* sesuai Algoritma 2.3.

5. Membandingkan desain *balanced sampling* dan desain Bayesian dengan desain sampling yang telah digunakan BPS dengan membandingkan nilai bias, *quality balancing*, varians, dan MSE sesuai kriteria desain sampling yang baik pada subbab 2.5.
6. Mendapatkan desain sampling terbaik dari ketiga desain yang dibandingkan.

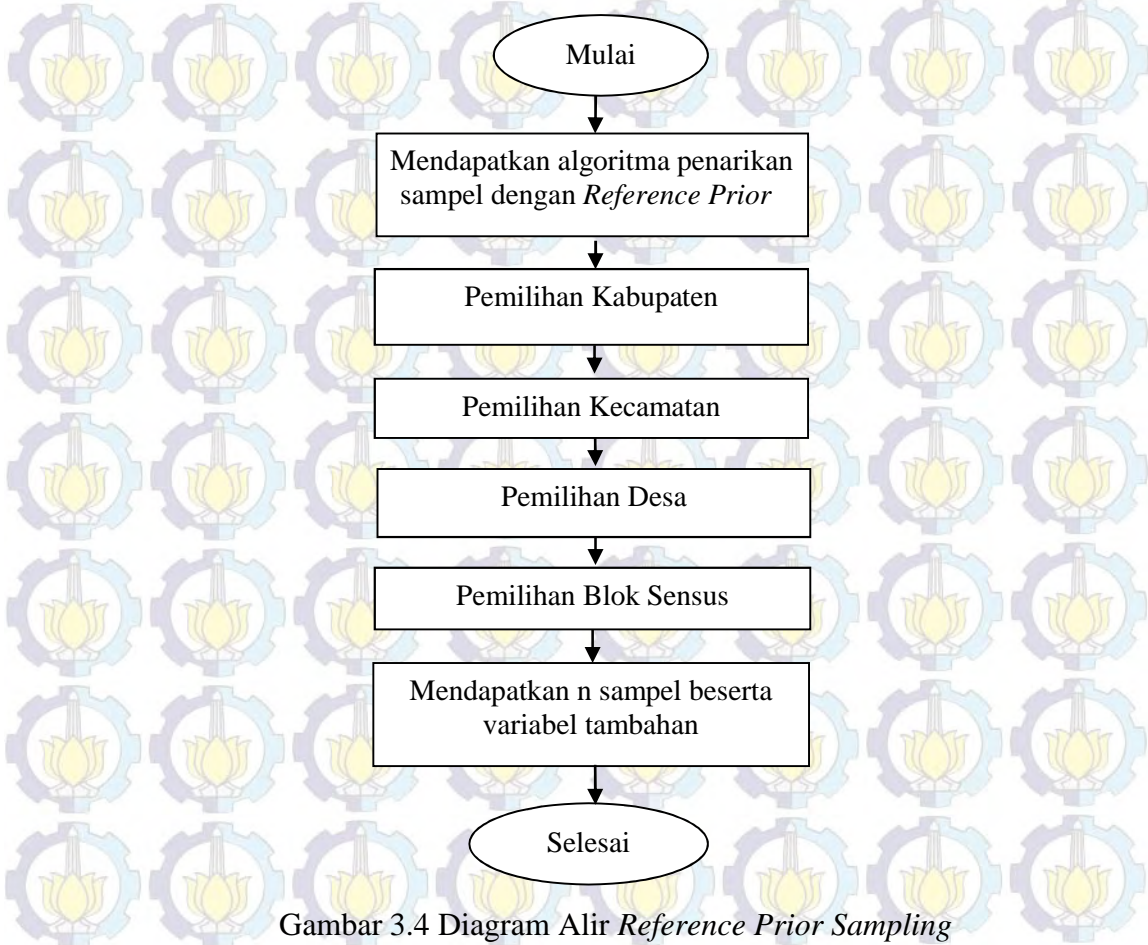
Langkah-langkah penelitian dapat dibentuk dalam diagram alir sebagaimana terdapat pada Gambar 3.2.



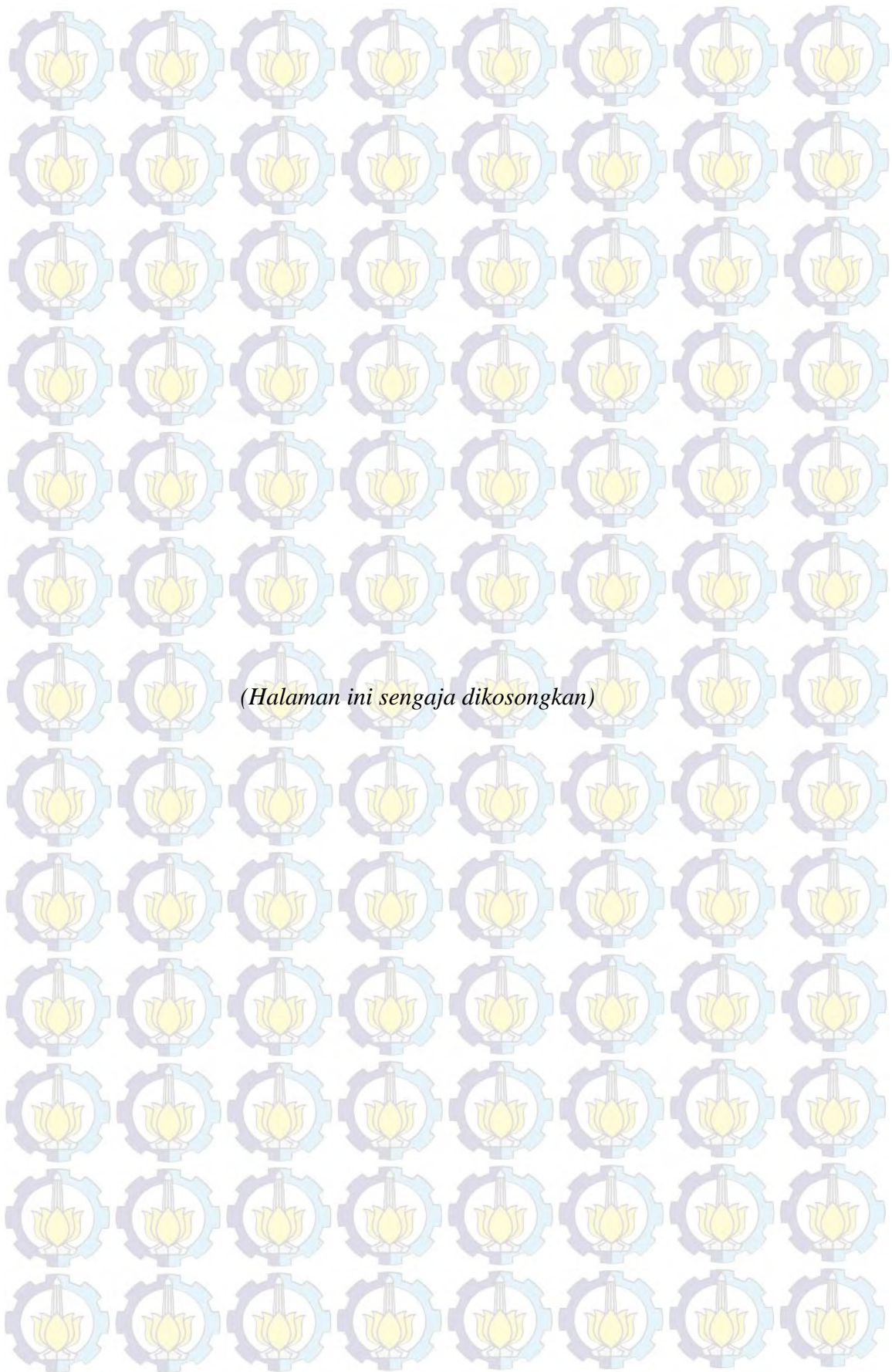
Gambar 3.2 Diagram Alir Penelitian



Gambar 3.3 Diagram Alir *Balanced Sampling* dengan *Cube Method*



Gambar 3.4 Diagram Alir *Reference Prior Sampling*



BAB 4

HASIL DAN PEMBAHASAN

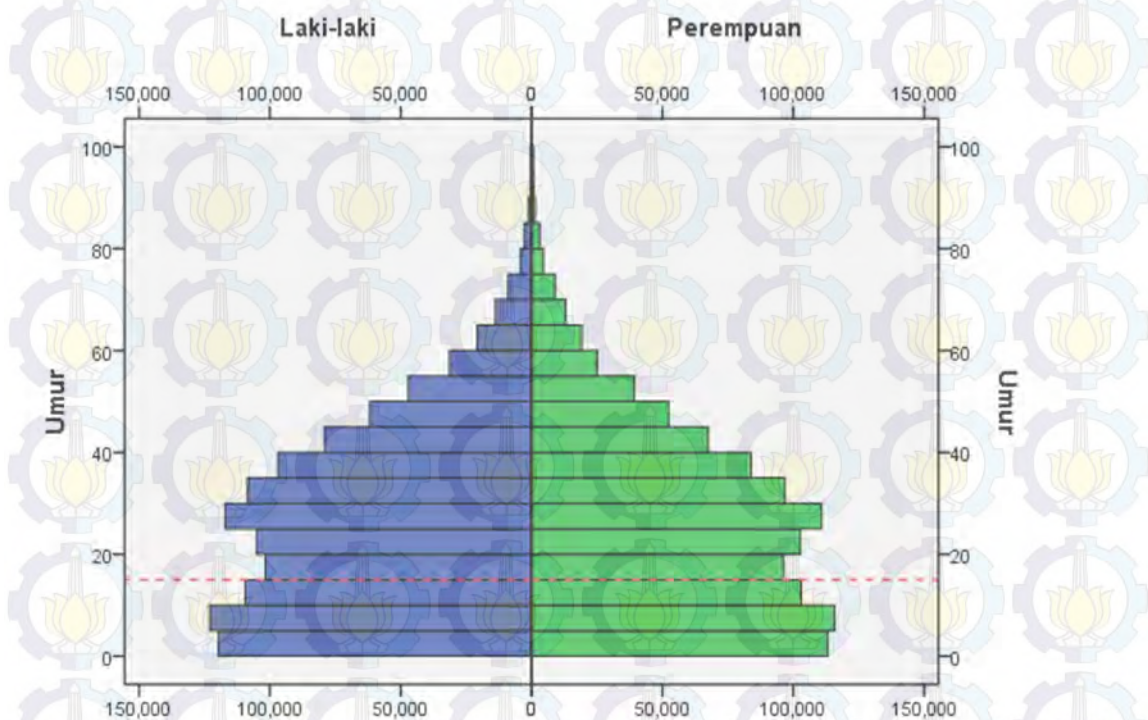
4.1 Gambaran Umum Provinsi Kalimantan Tengah

Provinsi Kalimantan Tengah terletak antara 0°45' Lintang Utara - 3° 30' Lintang Selatan dan 110°45' Bujur Timur - 115°51' Bujur Timur. Provinsi ini berada diantara tiga Provinsi yang lain yaitu Provinsi Kalimantan Barat, Provinsi Kalimantan Timur, dan Provinsi Kalimantan Selatan. Luas wilayah Kalimantan Tengah sebesar 153.927,81 Km² atau 8,04 persen dari total luas daratan Indonesia. Luas ini terbagi ke dalam 13 kabupaten dan 1 Kota, terdiri atas 130 kecamatan dan 1.528 desa/ kelurahan termasuk unit pemukiman transmigrasi (UPT).

Hasil Sensus Penduduk 2010 (SP2010) menunjukkan bahwa jumlah rumah tangga di Kalimantan Tengah sebanyak 572.842 rumah tangga dengan jumlah penduduk sebesar 2.212.089 jiwa. Jumlah penduduk ini terdiri dari 52 persen laki-laki dan 48 persen perempuan. Rasio jenis kelamin Kalimantan tengah sebesar 109, artinya setiap 100 penduduk perempuan terdapat 109 penduduk laki-laki. Kepadatan penduduk Kalimantan Tengah pada tahun 2010 hanya 14 jiwa per Km². Wilayah dengan kepadatan penduduk tertinggi adalah Kota Palangka Raya dengan 85 jiwa per Km², sedangkan wilayah dengan kepadatan terendah adalah Kabupaten Murung Raya dengan 4 jiwa per Km². Dari empat belas kabupaten/kota di Kalimantan Tengah, tiga kabupaten dengan persentase penduduk terbesar adalah Kabupaten Kotawaringin Timur (16,92 persen), Kabupaten Kapuas (14,90 persen), dan Kabupaten Kotawaringin Barat (10,66 persen).

Salah satu keunggulan data yang diperoleh dari SP2010 adalah diperolehnya data yang menyeluruh dan mendetail, misalnya diperolehnya data penduduk umur tunggal. Dari data tersebut dapat dibentuk penduduk menurut kelompok umur dan piramida penduduk. Piramida penduduk pada Gambar 4.1 menunjukkan bahwa penduduk Provinsi Kalimantan Tengah tergolong penduduk muda karena sebagian besar penduduk berada pada kelompok umur muda. Penduduk di bawah 15 tahun sebesar 31 persen, penduduk di antara 15 s.d 64 tahun sebesar 66 persen dan penduduk di atas 65 tahun sebesar 3 persen. Terkait

dengan hal tersebut, rasio ketergantungan penduduk sebesar 51,13 persen, dengan rincian rasio ketergantungan pada usia muda sebesar 46,75 persen dan rasio ketergantungan pada usia tua sebesar 4,39 persen. Pada kelompok umur 20 s.d 34 tahun, piramida menunjukkan grafik yang lebih lebar daripada kelompok umur dibawahnya. Hal ini diakibatkan banyaknya para pekerja yang masuk ke wilayah Kalimantan Tengah, misalnya untuk bekerja di sektor perkebunan maupun jasa pemerintah.



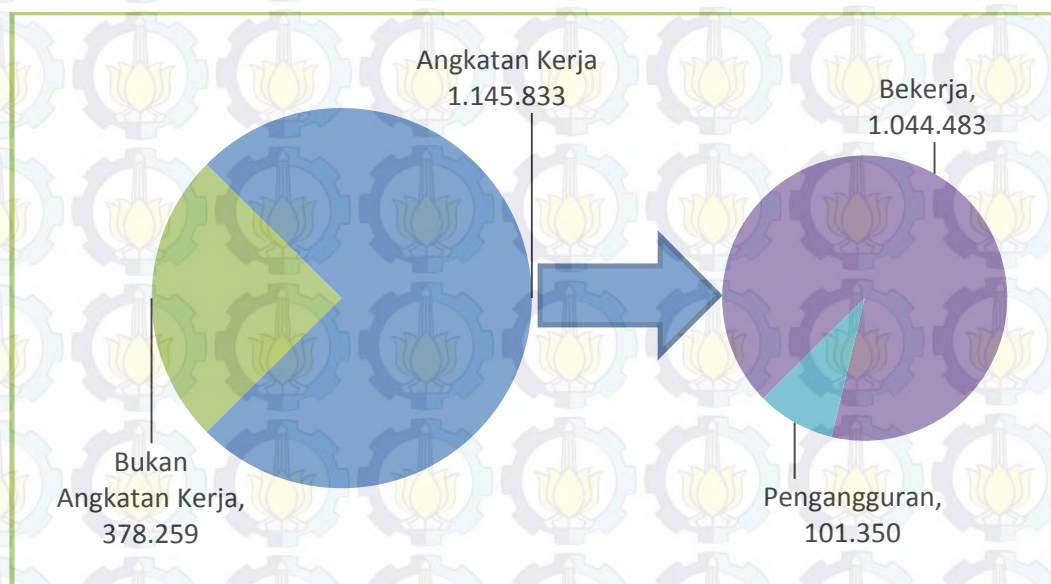
Gambar 4.1 Piramida Penduduk Kalimantan Tengah Tahun 2010

4.2 Gambaran Kondisi Ketenagakerjaan Provinsi Kalimantan Tengah Berdasarkan Hasil Sensus Penduduk 2010

Dalam Sensus Penduduk 2010, pertanyaan mengenai ketenagakerjaan hanya ditujukan kepada penduduk usia 15 tahun ke atas. Pertanyaan mengenai ketenagakerjaan hanya terdapat pada dokumen C1 yang digunakan untuk mencacah blok sensus biasa. Pada SP2010, blok sensus biasa di Kalimantan Tengah berjumlah 7.719 BS. Semua blok sensus ini yang akan menjadi kerangka

sampel dalam penarikan sampel dengan pendekatan Bayesian maupun dengan *balanced sampling*.

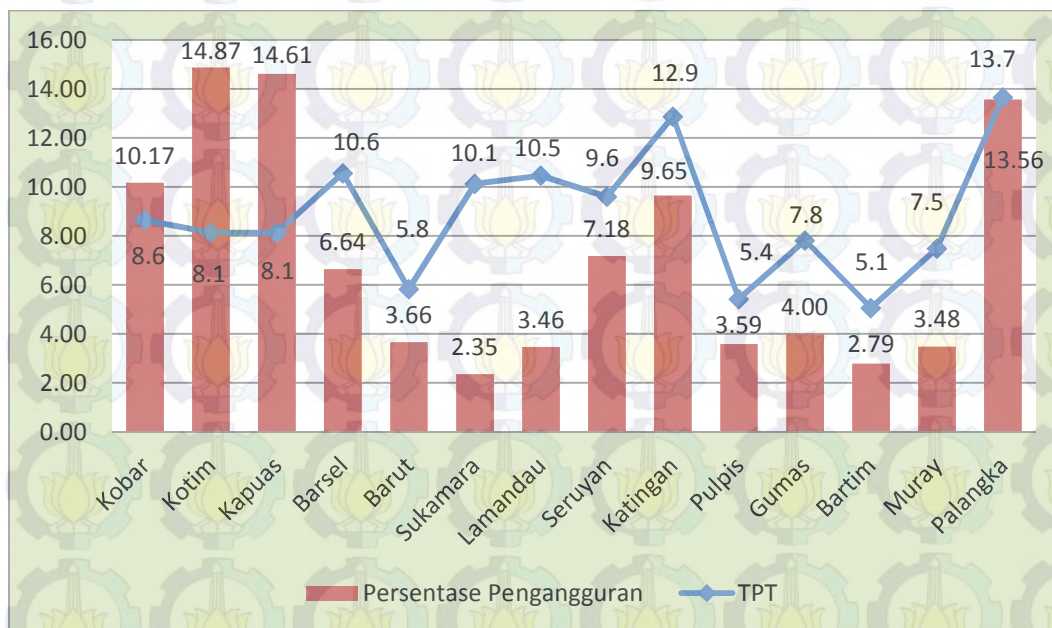
Penduduk usia kerja (15 tahun keatas) dibagi menjadi angkatan kerja dan bukan angkatan kerja. Jumlah penduduk usia kerja di Kalimantan Tengah sebesar 1.145.833 jiwa dengan persentase angkatan kerja sekitar 75 persen (Tingkat Partisipasi Angkatan Kerja/ TPAK) dan 25 persen bukan angkatan kerja yang terdiri dari ibu rumah tangga, anak sekolah dan lainnya. Dari total angkatan kerja, 91,15 persen di antaranya memiliki status bekerja dan 8,85 persen lainnya berstatus pengangguran. Persentase jumlah pengangguran terhadap total angkatan kerja ini menunjukkan Tingkat Pengangguran Terbuka (TPT). Rincian pembagian penduduk di atas 15 tahun ini dapat dilihat pada Gambar 4.2.



Gambar 4.2 Diagram Pembagian Penduduk Usia 15 Tahun ke Atas

Jumlah pengangguran sebesar 101.350 jiwa di Kalimantan Tengah tersebar ke 14 kabupaten/ kota. Kabupaten/ kota yang memiliki jumlah pengangguran terbesar adalah Kabupaten Kotawaringin Timur sejumlah 15.066 jiwa (14,87 persen dari provinsi), Kabupaten Kapuas sejumlah 14.808 jiwa (14,61 persen dari provinsi), dan Kota Palangka Raya sejumlah 13.744 jiwa (13,7 persen dari provinsi). Namun, tidak semua daerah yang memiliki jumlah pengangguran yang tinggi memiliki TPT yang tinggi pula. Pada Kabupaten Kotawaringin Timur dan

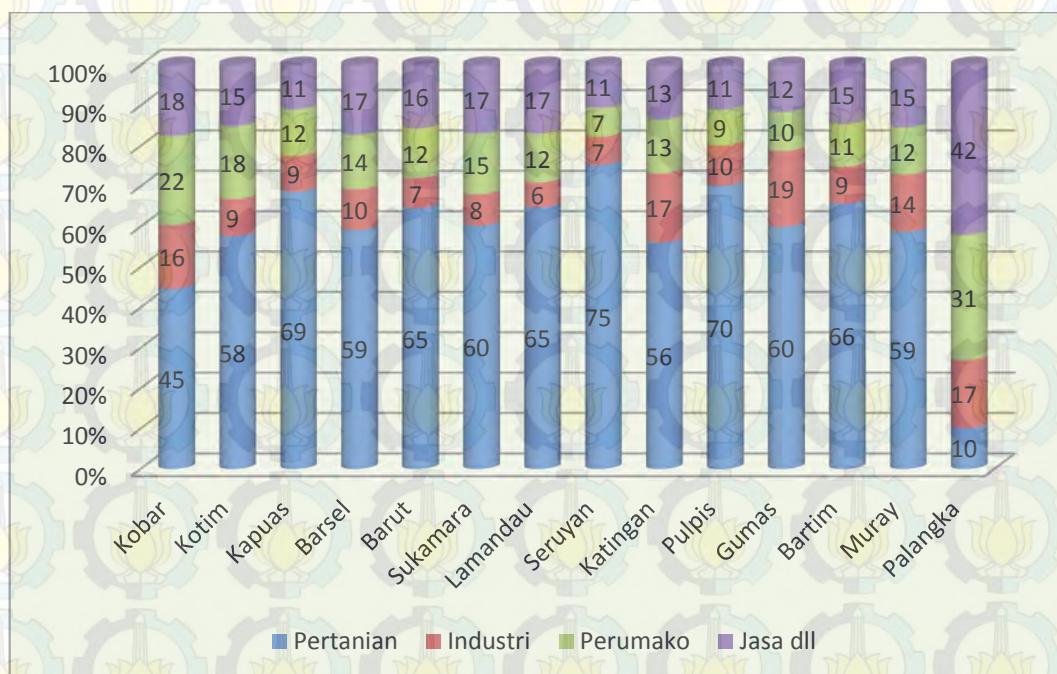
Kabupaten Kapuas, TPT tidak terlalu tinggi bila dibandingkan dengan kabupaten lainnya. Sebaliknya, Kabupaten Lamandau dan Kabupaten Sukamara yang memiliki jumlah pengangguran yang kecil (3,46 dan 2,35 persen dari provinsi) ternyata memiliki TPT yang cukup tinggi, yaitu 10,5 dan 10,1 persen. Hal ini dikarenakan perbedaan pembanding terhadap jumlah pengangguran. Perbandingan kedua nilai ini dapat dilihat pada Gambar 4.3.



Gambar 4.3 Persentase Pengangguran terhadap Provinsi dan TPT Kabupaten/ Kota di Kalimantan Tengah

Penduduk yang bekerja dapat dikelompokkan menurut lapangan usahanya. Secara sederhana, pengelompokan lapangan usaha dibagi menjadi sektor primer, sektor sekunder, dan sektor tersier. Pada Sensus Penduduk 2010 pertanyaan mengenai lapangan usaha utama dirinci hingga 19 sektor. Namun, pada penelitian ini hanya akan dijabarkan sampai dengan 4 kelompok. Hal ini mengikuti rincian pekerja menurut lapangan usaha yang disampaikan dalam Berita Resmi Statistik Ketenagakerjaan. Dengan pengelompokan tersebut, dari penduduk yang bekerja sebesar 1.044.483 jiwa, pekerja di sektor pertanian sebesar 57,24 persen, di sektor industri sebesar 11,33 persen, di sektor perdagangan, rumah makan, dan jasa akomodasi sebesar 15,11 persen serta di sektor jasa kemasyarakatan dan lainnya

sebesar 16,32 persen. Sebaran pekerja menurut lapangan usaha/ sektor juga dapat dilihat pada tingkat kabupaten. Gambar 4.4 menunjukkan bahwa sebagian besar pekerja bekerja pada sektor pertanian. Hal ini ditunjukkan dengan 13 kabupaten/ kota memiliki pekerja di sektor pertanian 45 persen atau lebih. Hal berbeda terjadi di kota Palangka Raya, dimana pekerja di sektor pertanian hanya 10 persen. Sebagai ibu kota provinsi, perekonomian di Palangka Raya lebih banyak menyerap pekerja di sektor jasa kemasyarakatan dan lainnya sebesar 42 persen serta di sektor perdagangan, perumahan, dan jasa akomodasi sebesar 31 persen. Selain Palangka Raya, daerah yang cukup besar persentase penyerapan pekerjaannya di kedua sektor ini Kabupaten Kotawaringi Barat, di sektor jasa kemasyarakatan dan lainnya sebesar 18 persen dan di sektor perdagangan, perumahan, dan jasa akomodasi sebesar 22 persen. Adapun daerah dengan persentase penyerapan pekerja di sektor industri yang cukup besar adalah Kabupaten Gunung mas (19 persen). Rincian lengkap persebaran pekerja menurut lapangan usaha di tiap kabupaten/ kota dapat dilihat pada Gambar 4.4.



Gambar 4.4 Persentase Penduduk Kalimantan Tengah yang Bekerja menurut Kabupaten/ Kota dan Lapangan Usaha

4.3 Karakteristik Data Ketenagakerjaan dari Sensus Penduduk 2010

Sebagaimana telah disampaikan pada subbab 2.8 dan subbab 4.2 bahwa pada individu yang tergolong angkatan kerja ada dua kemungkinan, yaitu bekerja atau menganggur. Bila dimisalkan individu yang menganggur merupakan kejadian sukses (1) dan individu yang tidak menganggur dikatakan kejadian gagal (0). Dengan demikian, maka tiap individu mengikuti distribusi Bernoulli dengan parameter misalkan PP . Bila beberapa individu angkatan kerja dikumpulkan di dalam satu wilayah, misalnya provinsi, kabupaten, kecamatan, desa, ataupun blok sensus, maka dapat dianalogikan terdapat beberapa kejadian menganggur atau tidak. Beberapa kejadian yang memiliki distribusi Bernoulli dengan parameter PP di suatu wilayah akan memiliki distribusi binomial dengan jumlah angkatan kerja sebanyak m dan proporsi pengangguran PP ($B(m, PP)$). Distribusi dari parameter binomial (PP) mempunyai distribusi *prior conjugate* distribusi beta-binomial (α, β). Dengan kerangka konsep atas data ini, pengangguran dianggap memiliki distribusi binomial. Dengan demikian, parameter proporsi pengangguran nantinya dapat diestimasi.

Tabel 4.1 Hasil Pengujian Distribusi terhadap Data Pengangguran pada Populasi dengan Uji Anderson Darling

H_0 = Mengikuti Distribusi:	Nilai Statistik Uji	Nilai Kritis ($\alpha=5\%$)	Kesimpulan
Uniform	2418,1	2,5018	Tolak H_0
Geometric	240,07	2,5018	Tolak H_0
Poisson	16376,0	2,5018	Tolak H_0
Bernoulli	-	-	
Binomial	-	-	
Hypergeometric	-	-	
Logarithmic	-	-	
Negative Binomial	-	-	

Selain itu, penggunaan kerangka konsep data digunakan karena dari hasil pengujian (*fitting*) distribusi terhadap data pengangguran di Kalimantan Tengah menunjukkan kesimpulan tidak mengikuti salah satu distribusi. Pengujian distribusi ini dilakukan terhadap data di 7.719 blok sensus. Tabel 4.1 merupakan ringkasan hasil pengujian *goodness of fit* dengan statistik uji Anderson Darling. Hasil pengujian menunjukkan bahwa H_0 ditolak yang artinya data pengangguran tidak memenuhi distribusi yang telah diuji sebagaimana pada Tabel 4.1. Karakteristik dari kerangka sampel penelitian ini dapat dilihat pada Tabel 4.2.

Tabel 4.2 Karakteristik Data Blok Sensus di Kerangka Sampel

Variabel	Total	Rata-rata	St. Dev	Max	Min
Angkatan Kerja	1.145.833	148,44	87,77	1.865	1
Pengangguran	101.350	13,13	17,55	212	0
Bekerja	1.044.483	135,31	83,12	1.833	1
Pertanian	597.862	77,45	81,44	1.667	0
Industri	118.288	15,32	20,96	256	0
Perdagangan, dll.	157.868	20,45	26,48	333	0
Jasa Kemasyarakatan, dll.	170.465	22,08	29,47	450	0

4.4 Hasil Penarikan Sampel

4.4.1 Karakteristik Sampel Sakernas 2011 dari Desain Sampling BPS

Sampel Sakernas triwulanan 2011 BPS Kalimantan Tengah ada sebanyak 116 blok sensus. Ringkasan karakteristik dari sampel dapat dilihat pada Tabel 4.3.

Tabel 4.3 Karakteristik Sampel dari Desain Sampling BPS

Variabel	Total	Rata-rata	St. Dev	Max	Min
Angkatan Kerja	20.359	175,51	77,13	668	60
Pengangguran	2.098	18,09	22,05	108	0
Bekerja	18.261	157,42	73,34	660	44
Pertanian	9.335	80,47	62,83	233	0
Industri	2.095	18,06	18,55	106	0
Perdagangan, dll.	3.068	26,45	27,13	116	0
Jasa Kemasyarakatan, dll.	3.763	32,44	51,48	450	0

Sampel Sakernas tidak mencakup variabel tambahan, tetapi variabel ditambahkan dalam rangka penelitian. Rincian sampel dapat dilihat pada Lampiran 2.

4.4.2 Penarikan Sampel *Balanced Sampling* dengan *Cube Method*

Penarikan sampel dengan *balanced sampling* menggunakan satu atau beberapa variabel yang digunakan sebagai variabel *balancing*. Dengan variabel ini sampel yang dipilih diupayakan menghasilkan estimasi dari variabel *balancing* mendekati nilai sebenarnya. Penelitian ini menggunakan tujuh variabel *balancing* yaitu jumlah angkatan kerja, jumlah pengangguran, jumlah yang bekerja, jumlah pekerja sektor pertanian, jumlah pekerja sektor industri, jumlah pekerja sektor perdagangan, rumah makan, dan jasa akomodasi, serta sektor jasa kemasyarakatan dan lainnya. Penarikan sampel dengan cara ini juga memerlukan suatu variabel yang menjadi penentu peluang terpilihnya sampel. Pada penelitian ini yang digunakan adalah jumlah penduduk dalam satu blok sensus. *Syntax* penarikan sampel dengan *cube method* dapat dilihat pada Lampiran 3 dan output *syntax* terdapat pada Lampiran 4

Dengan menggunakan jumlah sampel yang digunakan dalam Sakernas triwulanan di Kalimantan Tengah, dilakukan penarikan sampel sebanyak 116 blok sensus. Dari sampel ini kemudian dapat diperoleh juga data variabel penelitian yang bersesuaian dengan blok sensusnya. Seluruh sampel yang terpilih beserta variabel tambahannya dapat dilihat pada Lampiran 5. Beberapa karakteristik dari sampel yang dihasilkan dapat dilihat pada Tabel 4.4.

Tabel 4.4 Karakteristik Sampel dari *Balanced Sampling*

Variabel	Total	Rata-rata	St. Dev	Max	Min
Angkatan Kerja	20.318	175,16	70,24	450	47
Pengangguran	1.701	14,66	16,61	70	0
Bekerja	18.617	160,49	68,90	439	35
Pertanian	10.248	88,34	76,13	376	0
Industri	2.211	19,06	22,57	118	0
Perdagangan, dll.	2.982	25,71	30,28	170	0
Jasa Kemasyarakatan, dll.	3.176	27,38	32,90	173	0

4.4.3 Penarikan Sampel Pendekatan Bayesian dengan *Reference Prior*

Penarikan sampel dengan *reference prior* menggunakan distribusi suatu variabel yang merupakan variabel penting di dalam penelitian. Pada penelitian ini variabel yang digunakan adalah jumlah pengangguran. Selain itu, digunakan pula konsep hirarki. Hal ini dikarenakan dalam suatu daerah telah terdapat suatu hirarki, yaitu provinsi, kabupaten, kecamatan, dan desa. Di BPS struktur hirarki ini di tambah blok sensus dalam rangka menyeimbangkan beban kerja dari petugas. Proses penarikan sampel ini dapat dilihat pada Algoritma 4.1. Dari algoritma tersebut kemudian dibuat syntax pada program R sebagaimana pada Lampiran 6. Output dari syntax tersebut berupa sampel blok sensus dan dapat dilihat pada Lampiran 7.

Algoritma 4.1 *Reference Prior Sampling*

Step 1. Inisialisasi/ set nilai awal

- a) Vektor kosong kabupaten terpilih **sample_kab**
- b) Vektor kosong kecamatan terpilih **sample_kec**
- c) Vektor kosong desa terpilih **sample_desa**
- d) Vektor kosong blok sensus terpilih **sample_bs**
- e) Matriks **sample_gab** merupakan gabungan vektor **sample_kab**, **sample_kec**, **sample_desa**, dan **sample_bs**
- f) Inisial $n=116$ jumlah sampel yang akan diambil
- g) Inisial $d=1$

Step 2. Proses Pemilihan Sampel

- a) Untuk $d=1 \dots n$
- b) Pemilihan kabupaten
 - i. Membuat kumulatif jumlah pengangguran pada tiap kabupaten dalam Provinsi Kalimantan Tengah
 - ii. Indeks kabupaten $e=1$
 - iii. Membangkitkan U yang berdistribusi $U(1, \text{total kumulatif})$.

- iv. Jika $U \leq$ kumulatif kabupaten ke- e maka dilanjutkan ke kabupaten selanjutnya, jika tidak maka kabupaten ke- e tersebut terpilih sebagai sampel kabupaten ke- d .
- v. Mendapatkan **sample_kab**[d]
- c) Pemilihan kecamatan
 - i. Untuk kabupaten terpilih, dibuat kumulatif jumlah pengangguran tiap kecamatan.
 - ii. Indeks kecamatan $f=1$
 - iii. Membangkitkan V yang berdistribusi $U(1, \text{total kumulatif})$.
 - iv. Jika $V \leq$ kumulatif kecamatan ke- f maka dilanjutkan ke kecamatan selanjutnya, jika tidak maka kecamatan ke- f tersebut terpilih sebagai sampel kecamatan ke- d .
 - v. Mendapatkan **sample_kec**[d]
- d) Pemilihan desa
 - i. Untuk kecamatan terpilih, dibuat kumulatif jumlah pengangguran tiap desa.
 - ii. Indeks desa $g=1$
 - iii. Membangkitkan J yang berdistribusi $U(1, \text{total kumulatif})$.
 - iv. Jika $J \leq$ kumulatif desa ke- g maka dilanjutkan ke desa selanjutnya, jika tidak maka desa ke- g tersebut terpilih sebagai sampel desa ke- d .
 - v. Mendapatkan **sample_desa**[d]
- e) Pemilihan blok sensus
 - i. Untuk desa terpilih, di buat kumulatif jumlah pengangguran tiap blok sensus.
 - ii. Indeks blok sensus h
 - iii. Membangkitkan K yang berdistribusi $U(1, \text{total kumulatif})$.
 - iv. Jika $K \leq$ kumulatif blok sensus ke- h maka dilanjutkan ke blok sensus selanjutnya, jika tidak maka blok sensus ke- h tersebut terpilih sebagai sampel blok sensus ke- d .
 - v. Mendapatkan **sample_bs**[d]
- f) Kembali ke Step 2.a)

Step 3. Tahap Akhir: mendapatkan sejumlah n sampel dalam **sample_gab**.

Penarikan sampel dilakukan hingga mendapatkan sebanyak 116 blok sensus. Jumlah sampel ini sama dengan yang digunakan dalam Sakernas triwulanan di Kalimantan Tengah yang selanjutnya digunakan untuk estimasi parameter tingkat provinsi. Tabel 4.5 menunjukkan beberapa karakteristik dari 116 sampel yang dihasilkan. Seluruh sampel yang terpilih beserta variabel tambahannya dapat dilihat pada Lampiran 8.

Tabel 4.5 Karakteristik Sampel dari *Reference Prior*

Variabel	Total	Rata-rata	St. Dev	Max	Min
Angkatan Kerja	22.394	193,05	96,58	774	49
Pengangguran	4.338	37,40	32,06	212	2
Bekerja	18.056	155,66	78,44	562	34
Pertanian	8.307	71,61	68,87	492	0
Industri	3.030	26,12	26,42	139	0
Perdagangan, dll.	3.345	28,84	40,14	333	0
Jasa Kemasyarakatan, dll.	3.374	29,09	31,39	175	0

4.4.3 Perbandingan Karakteristik Sampel yang Dihasilkan

Pada tiga subbab di atas telah diperlihatkan bagaimana karakteristik sampel dari ketiga desain sampling. Sampel blok sensus yang dihasilkan tersebar ke 14 kabupaten/ kota yang ada di Kalimantan Tengah. Tabel 4.6 menunjukkan persebaran sampel beserta karakteristiknya. Dari sampel desain sampling BPS, proporsi pengangguran terhadap angkatan kerja sebesar 0,1031. Proporsi pekerja menurut lapangan usaha pertanian, industri, perdagangan, dan jasa sebesar 0,5112, 0,1147, 0,1680, dan 0,2061. Dari sampel *balanced sampling* dengan *cube method*, proporsi jumlah pengangguran terhadap jumlah angkatan kerja sebesar 0,0837. Proporsi pekerja menurut lapangan usaha pertanian, industri, perdagangan, dan jasa sebesar 0,5505, 0,1188, 0,1602, dan 0,1706. Adapun dari sampel *reference prior sampling*, proporsi jumlah pengangguran terhadap jumlah angkatan kerja sebesar 0,1937. Proporsi pekerja menurut lapangan usaha pertanian, industri, perdagangan, dan jasa sebesar 0,4601, 0,1678, 0,1853, dan 0,1869.

Tabel 4.6 Persebaran Sampel dan Karakteristik Sampel di 14 Kabupaten/ Kota

Desain	Kabu- paten	Jml BS Sampel	X1	X2	X3	X4	X5	X6	X7
BPS	6201	10	1848	206	1642	571	317	415	339
	6202	12	2317	334	1983	897	244	421	421
	6203	11	1541	65	1476	1156	45	108	167
	6204	8	1110	103	1007	427	105	192	283
	6205	8	1383	153	1230	697	98	203	232
	6206	5	780	69	711	340	108	141	122
	6207	7	1400	245	1155	623	54	209	269
	6208	8	1292	216	1076	633	108	142	193
	6209	9	1261	166	1095	517	196	181	201
	6210	8	1334	33	1301	729	174	115	283
	6211	7	1238	49	1189	919	78	96	96
	6212	7	1279	104	1175	840	96	128	111
	6213	7	1905	108	1797	697	180	338	582
	6271	9	1671	247	1424	289	292	379	464
	Jumlah	116	20359	2098	18261	9335	2095	3068	3763
	Proporsi			0,1031		0,5112	0,1147	0,1680	0,2061
Balanced Cube	6201	16	3111	256	2855	1386	269	686	514
	6202	24	4318	230	4088	2398	460	683	547
	6203	13	2012	128	1884	1166	197	263	258
	6204	7	939	133	806	513	140	89	64
	6205	4	734	27	707	576	11	50	70
	6206	4	517	36	481	270	41	61	109
	6207	1	91	8	83	68	1	2	12
	6208	10	1781	173	1608	1030	144	130	304
	6209	7	1177	165	1012	377	204	319	112
	6210	7	1237	60	1177	845	56	149	127
	6211	6	1206	131	1075	518	323	86	148
	6212	2	219	7	212	160	4	11	37
	6213	6	1288	107	1181	865	118	68	130
	6271	9	1688	240	1448	76	243	385	744
	Jumlah	116	20318	1701	18617	10248	2211	2982	3176
	Proporsi			0,0837		0,5505	0,1188	0,1602	0,1706
Reference	6201	11	2656	428	2228	1374	346	334	174
	6202	14	2722	606	2116	889	374	500	353
	6203	16	3232	677	2555	1123	400	694	338
	6204	7	1348	253	1095	575	168	133	219
	6205	3	592	112	480	220	23	89	148
	6206	2	486	103	383	144	59	91	89
	6207	4	873	234	639	360	56	95	128
	6208	6	990	168	822	359	159	133	171
	6209	14	2295	588	1707	722	387	218	380
	6210	10	1694	182	1512	1035	141	137	199
	6211	5	1083	205	878	492	243	60	83
	6212	2	378	12	366	234	40	49	43
	6213	5	1121	181	940	629	144	55	112
	6271	17	2924	589	2335	151	490	757	937
	Jumlah	116	22394	4338	18056	8307	3030	3345	3374
	Proporsi			0,1937		0,4601	0,1678	0,1853	0,1869

4.5 Perbandingan Hasil Estimasi dari Ketiga Desain Sampling

Dua hal yang akan dijadikan objek perbandingan dari ketiga desain adalah proporsi pengangguran serta proporsi pekerja menurut lapangan usaha. Dengan demikian, diperlukan sarana yang dapat dijadikan pembanding. Hal pertama yang dilakukan adalah melakukan estimasi dari sampel ketiga desain. Desain BPS dan *balanced* sampling dilakukan dua cara estimasi, yaitu estimasi jumlah total variabel dengan *design based* dan estimasi parameter distribusi dengan metode Bayesian. Adapun pada *reference prior sampling* hanya dilakukan estimasi parameter distribusi dengan metode Bayesian. Hasil estimasi akan digunakan pada langkah selanjutnya, yaitu perbandingan di antara ketiga desain.

Pada sampel Sakernas 2011/ sampel dari desain sampling BPS, hal pertama yang dilakukan adalah mengestimasi total dari tujuh variabel penelitian, yaitu angkatan kerja, pengangguran, bekerja, pekerja di sektor pertanian, pekerja di sektor industri, pekerja di sektor perdagangan, rumah makan dan akomodasi serta pekerja di sektor jasa kemasyarakatan dan lainnya. Kemudian dihitung proporsi pengangguran, proporsi bekerja dan proporsi pekerja menurut lapangan usahanya. Hasil estimasi dari desain sampling BPS sebagaimana pada Tabel 4.7.

Tabel 4.7 Hasil Estimasi Jumlah Nilai Karakteristik dari Sampel Desain Sampling BPS dengan Metode *Design Based*

Variabel	Estimasi Total	Proporsi	Keterangan
Angkatan Kerja	1.159.308		
Pengangguran	122.015	0,1052	Proporsi terhadap jumlah angkatan kerja
Bekerja	1.037.293	0,8948	
Pertanian	604.877	0,5831	Proporsi terhadap jumlah angkatan kerja yang bekerja (pekerja)
Industri	115.407	0,1113	
Perdagangan, dll.	149.830	0,1444	
Jasa Kemasyarakatan, dll.	167.180	0,1612	

Adapun hasil estimasi Horvitz-Thompson pada jumlah dan nilai proporsi variabel dari *balanced sampling* sesuai dengan persamaan (2.10) sebagaimana pada Tabel 4.8:

Tabel 4.8 Hasil Estimasi Jumlah Nilai Karakteristik dari Sampel *Balanced Sampling* dengan Metode Horvitz Thompson

Variabel	Estimasi jumlah	Proporsi	Keterangan
Angkatan Kerja	1.148.324		
Pengangguran	101.474	0,0884	Proporsi terhadap jumlah angkatan kerja
Bekerja	1.046.851	0,9116	
Pertanian	601.015	0,5741	Proporsi terhadap jumlah angkatan kerja yang bekerja (pekerja)
Industri	117.986	0,1127	
Perdagangan, dll.	159.881	0,1527	
Jasa Kemasyarakatan, dll.	167.969	0,1605	

Selanjutnya dilakukan estimasi proporsi pengangguran dan proporsi pekerja menurut lapangan usaha dengan metode Bayesian. *Syntax* dan hasil estimasi parameter distribusi binomial dan multinomial dengan metode Bayesian dari sampel dengan desain BPS, *balanced sampling*, dan *reference prior sampling* dapat dilihat pada Lampiran 9. sampai dengan Lampiran 17. Ringkasan hasil yang menunjukkan proporsi pengangguran dan proporsi pekerja menurut lapangan usaha ini akan ditunjukkan pada subbab selanjutnya.

4.5.1 Perbandingan Hasil Estimasi Proporsi Pengangguran

Perbandingan proporsi pengangguran yang pertama dengan menggunakan hasil estimasi metode *design based* pada desain BPS dan *balanced sampling* serta estimasi metode Bayesian pada *reference prior sampling*. Nilai proporsi pengangguran yang terdapat pada Tabel 4.7 untuk desain sampling BPS dan pada Tabel 4.8 untuk *balanced sampling* diambil untuk dibandingkan dengan nilai estimasi proporsi pengangguran dari *reference prior sampling* dengan MCMC dan *Gibbs Sampler* pada Tabel 4.9.

Tabel 4.9 Perbandingan Estimasi Proporsi Pengangguran dengan Metode Estimasi Berbeda

Desain	Estimasi Proporsi	Parameter Proporsi	Bias	RD	Varians	MSE
BPS	0,1052	0,0885	0,01680	18,99	1,96E-05	3,02E-04
<i>Balanced</i>	0,0884	0,0885	0,00008	0,10	3,04E-06	3,05E-06
<i>Reference</i>	0,0902	0,0885	0,00177	2,00	6,96E-08	3,20E-06

Tabel 4.9 menunjukkan nilai estimasi proporsi ketiga desain dengan nilai proporsi populasi. Desain sampling BPS menghasilkan nilai absolut bias sebesar 0,01680 dan nilai absolut *relative deviation* sebesar 18,99 persen. *Balanced sampling* menghasilkan nilai absolut bias sebesar 0,00008 dan nilai absolut *relative deviation* sebesar 0.095 persen. Adapun *reference prior sampling* menghasilkan nilai bias sebesar 0,00177 dan nilai *relative deviation* sebesar 2 persen. Jadi, *balanced sampling* memberikan tingkat validitas yang lebih baik untuk proporsi pengangguran dengan metode estimasi yang berbeda. Bila dilihat nilai varians-nya, estimasi proporsi pengangguran dari sampel hasil *reference prior sampling* dengan metode Bayesian menghasilkan varians terkecil. Selanjutnya, perbandingan nilai MSE menunjukkan *balanced sampling* menghasilkan nilai terkecil, yaitu sebesar $3,05 \times 10^{-6}$. Dengan demikian, bila menggunakan metode estimasi yang berbeda, *balanced sampling* memberikan sampel yang menghasilkan estimasi yang lebih akurat.

Perbandingan proporsi pengangguran yang kedua dilakukan dengan menggunakan nilai estimasi parameter distribusi dari variabel pengangguran. Dengan jumlah m jumlah angkatan kerja dan PP adalah proporsi pengangguran terhadap angkatan kerja. Dengan menggunakan prior *conjugate* dari distribusi binomial yaitu beta-binomial dan informasi dari populasi yang dimiliki (hasil SP2010), didapatkan estimasi dari parameter PP (proporsi pengangguran terhadap angkatan kerja). Perbandingan kedua dilakukan dengan mengestimasi semua sampel ketiga desain dengan metode Bayesian, yaitu MCMC dan *Gibbs Sampler*, hasil estimasi dapat dilihat pada Tabel 4.10.

Tabel 4.10 Perbandingan Estimasi Proporsi Pengangguran dengan Metode Estimasi Bayesian

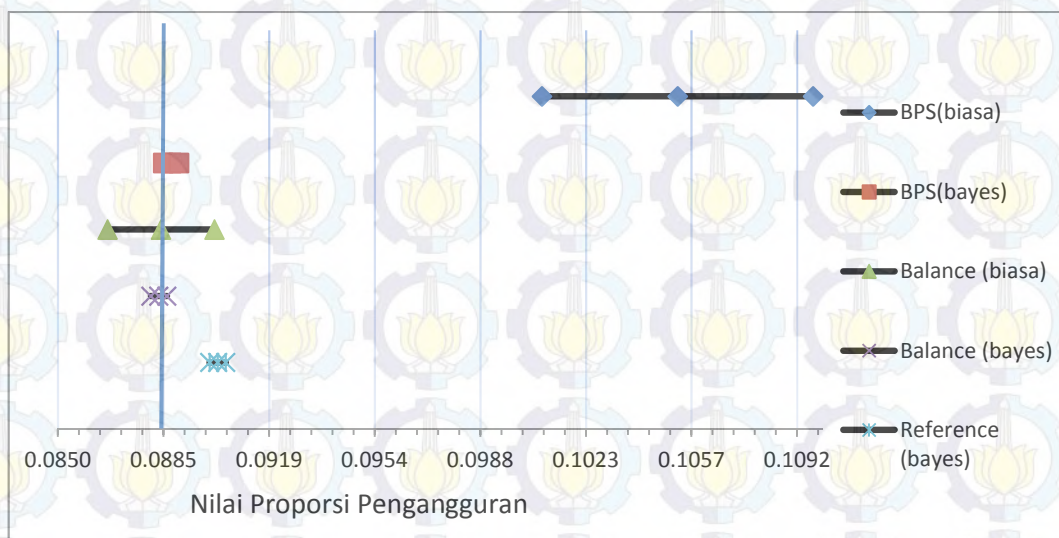
Desain	Estimasi Proporsi	Parameter Proporsi	Bias	RD	Varians	MSE
BPS	0,0887	0,0885	0,00025	0,28	6,67E-08	1,29E-07
<i>Balanced</i>	0,0884	0,0885	0,00009	0,10	6,83E-08	7,66E-08
<i>Reference</i>	0,0902	0,0885	0,00177	2,00	6,96E-08	3,20E-06

Keterangan: prior beta(101350,1044483)

Dari proses MCMC dan *Gibbs Sampler* diperoleh nilai estimasi proporsi dan nilai standar deviasi. Selanjutnya dihitung nilai bias, *relative deviation*, *varians*, dan MSE. Desain *balanced sampling* menghasilkan nilai bias, *relative deviation*, dan MSE terkecil, yaitu sebesar 0,00009, 0,10 persen, dan $7,66 \times 10^{-8}$. Dengan demikian, *balanced sampling* memberikan tingkat validitas dan akurasi yang lebih baik dibandingkan dengan desain lainnya. Adapun tingkat reliabilitas desain sampling BPS lebih baik dari kedua desain lain karena nilai *varians* yang dihasilkan merupakan yang terkecil dengan nilai sebesar $6,67 \times 10^{-8}$.

Dengan mempertimbangkan perbandingan dengan menggunakan metode estimasi yang sama dan perbandingan dengan menggunakan metode estimasi yang berbeda, *balanced sampling* dengan *cube method* merupakan desain sampling yang lebih baik dalam memberikan sampel untuk mengestimasi nilai proporsi pengangguran.

Perbandingan antar desain sampling juga dapat dilakukan dengan melihat gambar/ grafik. Dari gambar dapat diketahui hasil estimasi yang relatif sama dan dapat diketahui hasil estimasi yang dapat menangkap nilai proporsi pengangguran yang sebenarnya. Gambar 4.5 menunjukkan bahwa hasil estimasi dari sampel BPS dengan metode estimasi biasa (*design based*) dapat dikatakan berbeda dengan hasil estimasi lainnya. Hal ini dikarenakan grafik nilai estimasi dengan penambahan dan pengurangan standar deviasi tidak beririsan/ sejajar dengan grafik hasil estimasi desain yang lain. Grafik *reference prior sampling* hanya beririsan dengan grafik *balanced sampling* dengan metode estimasi biasa (HT) sehingga dapat dikatakan terdapat kesamaan hasil estimasi.



Gambar 4.5 Perbandingan Nilai Estimasi Proporsi Pengangguran dengan Pengurangan dan Penambahan Standar Deviasi

Gambar 4.5 juga menunjukkan bahwa terdapat tiga hasil estimasi yang memiliki kesamaan, yaitu desain sampling BPS dengan metode estimasi Bayesian, desain *balanced sampling* dengan metode estimasi biasa (HT), dan *balanced sampling* dengan metode estimasi Bayesian. Grafik estimasi dengan penambahan dan pengurangan standar deviasi dari ketiganya juga beririsan/ sejajar dengan nilai proporsi pengangguran yang sebenarnya, yaitu 0,0845. Bila menggunakan cara perbandingan tersebut, maka ketiga desain beserta cara estimasinya lebih diunggulkan. Kemudian dapat dipilih yang lebih mudah dalam pelaksanaannya.

4.5.2 Perbandingan Hasil Estimasi Proporsi Pekerja menurut Lapangan Usaha

Sebagaimana perbandingan pada proporsi pengangguran, perbandingan pada proporsi pekerja menurut lapangan usaha dilakukan dua kali. Pertama dengan metode estimasi yang berbeda dan kedua dengan metode estimasi yang sama yaitu MCMC dan *Gibbs Sampler* pada semua desain. Perbandingan pertama dilakukan dengan mengambil nilai proporsi pekerja menurut lapangan usaha yang terdapat pada Tabel 4.7 untuk desain sampling BPS dan nilai proporsi pekerja menurut lapangan usaha pada Tabel 4.8 untuk *balanced sampling*.

Tabel 4.11 Perbandingan Estimasi Proporsi Proporsi Pekerja menurut Lapangan Usaha dengan Metode Estimasi Berbeda

Desain		Estimasi Proporsi	Parameter Proporsi	Bias	RD	Varians	MSE
BPS	p1	0,5831	0,5724	0,01073	1,8745	2,58E-05	1,41E-04
	p2	0,1113	0,1133	0,00199	1,7594	8,53E-06	1,25E-05
	p3	0,1444	0,1511	0,00670	4,4340	7,68E-06	5,26E-05
	p4	0,1612	0,1632	0,00204	1,2473	1,07E-05	1,48E-05
<i>Balance</i>	p1	0,5741	0,5724	0,00172	0,3001	1,15E-06	4,10E-06
	p2	0,1127	0,1133	0,00054	0,4810	3,90E-06	4,20E-06
	p3	0,1527	0,1511	0,00158	1,0460	4,03E-06	6,53E-06
	p4	0,1605	0,1632	0,00275	1,6874	4,05E-06	1,16E-05
<i>Reference</i>	p1	0,5670	0,5724	0,00540	0,9434	2,29E-07	2,94E-05
	p2	0,1160	0,1133	0,00275	2,4280	9,66E-08	7,66E-06
	p3	0,1525	0,1511	0,00136	0,8967	1,22E-07	1,96E-06
	p4	0,1645	0,1632	0,00129	0,7934	1,28E-07	1,80E-06

Perbandingan hasil estimasi ketiga desain sampling melibatkan nilai populasi proporsi pekerja menurut lapangan usaha. Setelah dikurangi nilai populasi/ nilai sebenarnya didapatkan nilai bias dari masing-masing proporsi. Pada proporsi pekerja di sektor pertanian dan sektor industri, nilai absolut dari bias desain *balanced sampling* merupakan nilai terkecil, yaitu sebesar 0,00172 dan 0,00054. Adapun pada proporsi pekerja di sektor perdagangan, rumah makan, dan jasa akomodasi serta sektor jasa kemasyarakatan dan lainnya, nilai absolut bias desain *reference prior sampling* merupakan nilai yang terkecil, yaitu sebesar 0,00136 dan 0,00129. Desain *balanced sampling* dan desain *reference prior sampling* lebih baik dibandingkan dari desain BPS bila dilihat dari sisi validitasnya.

Bila dilihat dari nilai varians yang dihasilkan, varians yang dihasilkan desain *reference prior sampling* dengan metode estimasi Bayesian menghasilkan nilai yang paling kecil pada keempat proporsi lapangan usaha. Dapat dikatakan bahwa *reference prior sampling* lebih reliable dari kedua desain yang lain.

Bila perbandingan dilakukan pada nilai MSE yang dihasilkan, pada proporsi pekerja di sektor pertanian dan sektor industri, nilai MSE desain *balanced sampling* menghasilkan nilai terkecil, yaitu sebesar $4,10 \times 10^{-6}$ dan $4,20 \times 10^{-6}$. Adapun pada proporsi pekerja di sektor perdagangan, rumah makan, dan jasa akomodasi serta sektor jasa kemasyarakatan dan lainnya, nilai MSE

desain *reference prior sampling* merupakan nilai yang terkecil, yaitu sebesar $1,96 \times 10^{-6}$ dan $1,80 \times 10^{-6}$. Namun, hasil penjumlahan MSE *balanced sampling* lebih kecil bila dibandingkan dengan hasil penjumlahan *reference prior sampling*, yaitu $2,65 \times 10^{-5}$ dibandingkan dengan $4,08 \times 10^{-5}$.

Tabel 4.12 Perbandingan Estimasi Proporsi Pekerja menurut Lapangan Usaha dengan Estimasi Metode Estimasi Bayesian

Desain		Estimasi Proporsi	Parameter Proporsi	Bias	RD	Varians	MSE
BPS	p1	0,5673	0,5724	0,0051	0,8910	2,29E-07	2,62E-05
	p2	0,1156	0,1133	0,0023	2,0748	9,63E-08	5,62E-06
	p3	0,1525	0,1511	0,0014	0,8967	1,22E-07	1,96E-06
	p4	0,1645	0,1632	0,0013	0,7934	1,28E-07	1,80E-06
Balance	p1	0,5668	0,5724	0,0056	0,9783	2,29E-07	3,16E-05
	p2	0,1157	0,1133	0,0024	2,1631	9,63E-08	6,10E-06
	p3	0,1527	0,1511	0,0016	1,0291	1,22E-07	2,54E-06
	p4	0,1648	0,1632	0,0016	0,9772	1,28E-07	2,67E-06
Reference	p1	0,5670	0,5724	0,0054	0,9434	2,29E-07	2,94E-05
	p2	0,1160	0,1133	0,0027	2,4280	9,66E-08	7,66E-06
	p3	0,1525	0,1511	0,0014	0,8967	1,22E-07	1,96E-06
	p4	0,1645	0,1632	0,0013	0,7934	1,28E-07	1,80E-06

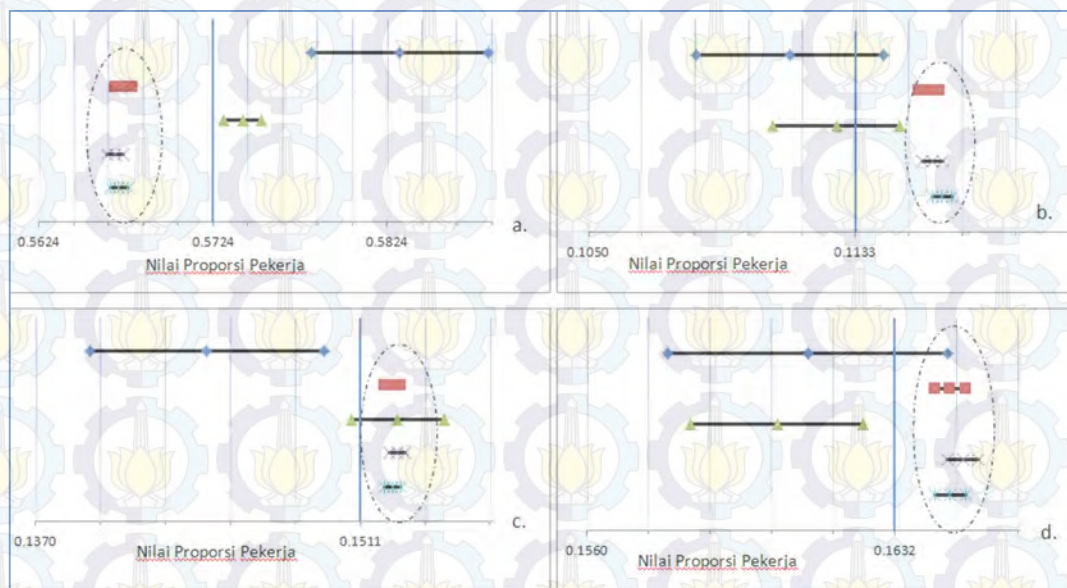
Keterangan: prior dirch(597862,118288,157868,170465)

Perbandingan proporsi pekerja menurut lapangan usaha yang kedua menggunakan metode estimasi Bayesian pada semua desain sampling. Tabel 4.12 memberikan rincian nilai perbandingan ketiga desain sampling. Metode estimasi Bayesian dengan MCMC dan *Gibbs Sampler* menghasilkan nilai estimasi proporsi dari keempat lapangan usaha dengan nilai standar deviasinya. Selanjutnya dihitung nilai bias, *relative deviation*, varians, dan MSE. Bila dilihat nilai bias yang dihasilkan, desain BPS memberikan tingkat validitas yang lebih baik dibandingkan desain lainnya pada sektor pertama dan kedua. Pada sektor ketiga dan keempat, *reference prior sampling* memberikan tingkat validitas yang sama dengan desain BPS. Dari nilai varians, ketiga desain memiliki nilai yang hampir sama. Dari nilai MSE, desain BPS memberikan tingkat akurasi yang lebih baik dibandingkan desain lainnya pada sektor pertama dan kedua. Pada sektor ketiga dan keempat, tingkat akurasi yang sama dengan *reference prior sampling*.

Dua perbandingan dengan metode estimasi sama dan metode estimasi berbeda menghasilkan dua desain yang unggul. Pada perbandingan dengan metode estimasi yang berbeda, *balanced sampling* dengan Horvitz-Thompson lebih unggul dari desain yang lain. Pada perbandingan dengan metode estimasi yang sama (metode Bayesian), desain BPS lebih unggul. Bila keduanya dibandingkan, nilai total MSE *balanced sampling* lebih kecil dari desain BPS dengan total MSE sebesar $2,65 \times 10^{-5}$ dibandingkan dengan $3,56 \times 10^{-5}$.

Selanjutnya dilakukan perbandingan estimasi proporsi pekerja menurut lapangan usaha dengan grafik. Gambar 4.6 menunjukkan empat grafik estimasi untuk sektor pertanian, sektor industri, sektor perdagangan, rumah makan, dan akomodasi, serta sektor jasa kemasyarakatan dan lainnya. Dari keempat grafik menunjukkan bahwa hasil estimasi dengan metode Bayesian pada ketiga desain sampling menghasilkan nilai yang relatif sama karena grafik beririsan/ sejajar. Hasil estimasi dengan metode Bayesian memberikan hasil yang relatif sama dengan hasil estimasi biasa *balanced sampling* pada sektor perdagangan dan dengan hasil estimasi biasa desain BPS pada sektor jasa.

Bila grafik estimasi dengan penambahan dan pengurangan standar deviasi di bandingkan dengan garis nilai proporsi sebenarnya, desain sampling dengan metode estimasi Bayesian tidak memotong garis tersebut. Adapun grafik desain BPS dan *balanced sampling* dengan estimasi biasa memotong garis tersebut masing-masing pada 2 sektor. Namun, pada grafik estimasi proporsi pekerja menurut lapangan usaha ini, faktor standar deviasi yang besar cukup mempengaruhi terpotongnya grafik oleh garis nilai proporsi sebenarnya. Sebagai contoh, pada Gambar 4.6 d. grafik desain BPS dan *balanced sampling* dengan estimasi biasa memotong garis nilai sebenarnya, tetapi nilai estimasi titik dari keduanya tidak lebih dari ketiga desain sampling yang menggunakan estimasi Bayesian.

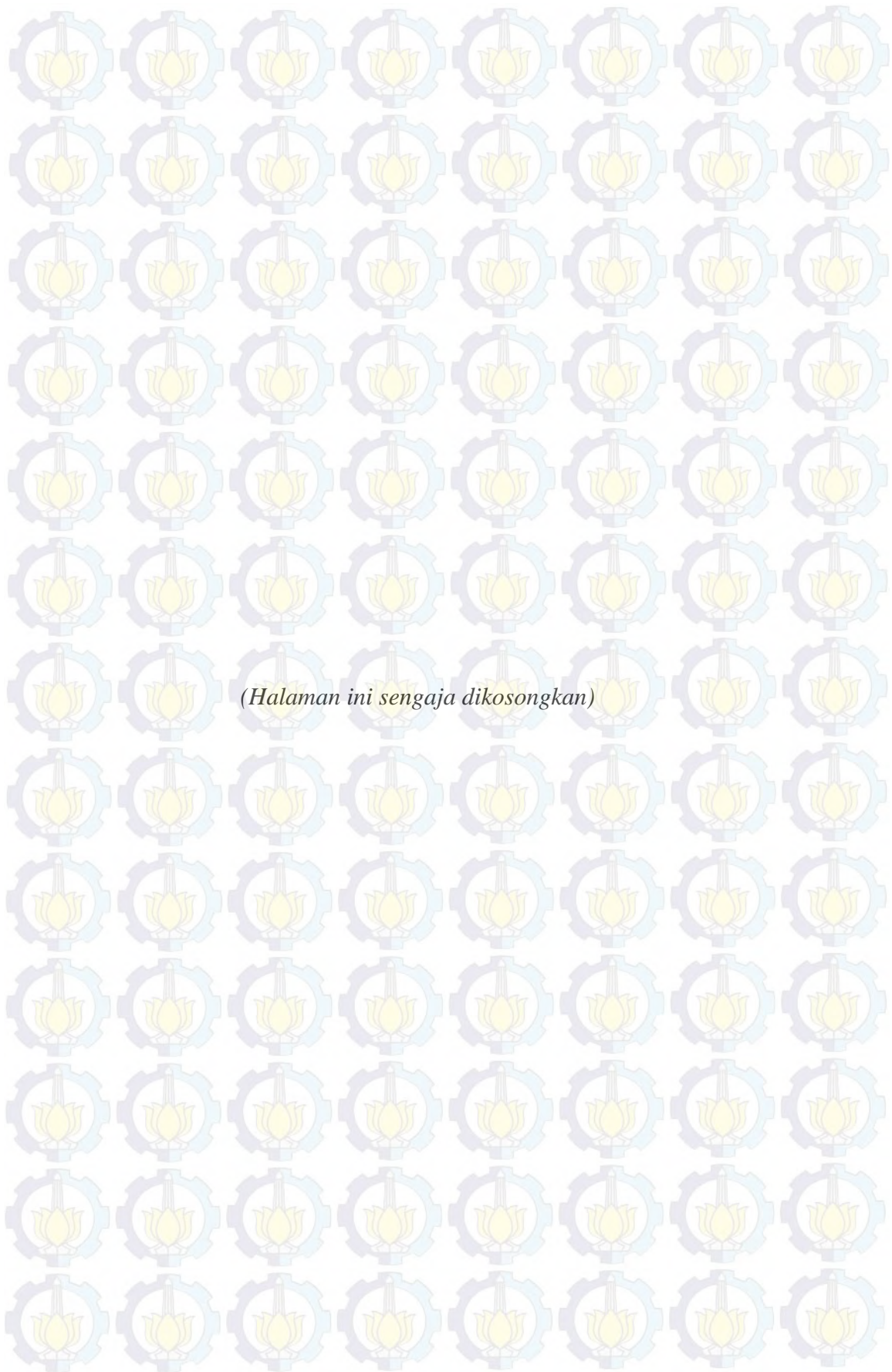


Gambar 4.6 Perbandingan Nilai Estimasi Proporsi Pekerja menurut Lapangan Usaha dengan Penambahan dan Pengurangan Standar Deviasi: a. Pertanian, b. Industri, c. Perdagangan, rumah makan dan akomodasi, d. Jasa Kemasyarakatan dll.

Keterangan:

- BPS(estimasi biasa)
- BPS(estimasi bayes)
- ▲— Balance (estimasi biasa)
- ×— Balance (estimasi bayes)
- *— Reference (estimasi bayes)

Dari dua perbandingan pada estimasi proporsi pengangguran dan proporsi pekerja menurut lapangan usaha, dengan tabel maupun grafik, dapat disimpulkan bahwa desain *balanced sampling* dengan *cube method* merupakan desain sampling yang lebih baik daripada desain sampling BPS dan *reference prior sampling*.



BAB 5

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dari kegiatan penelitian yang telah dilakukan dan telah diuraikan pada bab sebelumnya, diperoleh beberapa kesimpulan sebagai berikut:

1. Dari sampel triwulanan Sakernas 2011 sebesar 116 blok sensus, dengan metode estimasi dengan *design based*, diperoleh nilai proporsi pengangguran sebesar 0,1052, absolut bias sebesar 0,0168, varians sebesar $1,96 \times 10^{-5}$, dan MSE sebesar $3,02 \times 10^{-4}$. Dengan metode Bayesian, diperoleh proporsi pengangguran sebesar 0,0887, absolut bias sebesar 25×10^{-5} , varians sebesar $6,67 \times 10^{-8}$, dan MSE sebesar $1,29 \times 10^{-7}$.
2. Pada penerapan *balanced sampling* dengan *cube method*, diperlukan beberapa kali penarikan sampel hingga mendapatkan sampel sebesar 116. Dengan metode estimasi Horvitz-Thompson, diperoleh nilai proporsi pengangguran sebesar 0,0884, absolut bias sebesar 8×10^{-5} , varians sebesar $3,04 \times 10^{-6}$, dan MSE sebesar $3,05 \times 10^{-6}$. Dengan metode Bayesian, diperoleh proporsi pengangguran sebesar 0,0883, absolut bias sebesar 15×10^{-5} , varians sebesar $6,85 \times 10^{-8}$, dan MSE sebesar $9,13 \times 10^{-8}$.
3. Penarikan sampling dengan *reference prior sampling* menggunakan distribusi *reference prior* dari data sebagai faktor terpilihnya suatu daerah, menggunakan prinsip hirarki, dan memperoleh sampel secara satu per satu hingga jumlah sampel yang ditentukan, yaitu 116. Setelah sampel yang diperoleh diestimasi dengan metode Bayesian, diperoleh estimasi parameter dari distribusi binomial sebesar 0,0902, artinya proporsi pengangguran terhadap angkatan kerja sebesar 0,0902. Selain itu, diperoleh absolut bias sebesar 0,02, varians sebesar $6,96 \times 10^{-8}$, dan MSE sebesar $3,20 \times 10^{-6}$.
4. Pada perbandingan estimasi proporsi pengangguran, sampel dari *balanced sampling* dengan *cube method* memberikan tingkat akurasi estimasi yang paling baik. Pada perbandingan estimasi proporsi pekerja menurut

lapangan usaha, pada perbandingan dengan menggunakan metode estimasi yang berbeda, sampel dari *balanced sampling* dengan *cube method* memberikan tingkat akurasi yang lebih baik dari desain lain. Adapun pada perbandingan dengan ketiga desain diestimasi dengan metode Bayesian, sampel dari desain BPS memberikan tingkat akurasi yang lebih baik dibandingkan desain lainnya. Dengan mempertimbangkan dua perbandingan tersebut, dapat disimpulkan bahwa *balanced sampling* dengan *cube method* memberikan sampel yang lebih baik untuk estimasi proporsi pengangguran dan estimasi proporsi pekerja menurut lapangan usaha.

5. Estimasi sampel dari desain sampling BPS dengan menggunakan metode Bayesian akan menghasilkan estimator yang lebih baik jika dibandingkan dengan metode estimasi berdasarkan *design based*.

5.2 Saran

Berdasarkan penelitian yang telah dilakukan, dapat direkomendasikan beberapa saran sebagai berikut:

1. Pada penelitian *balanced sampling* dengan *cube method*, dapat diteliti bagaimana dampak dari kombinasi *balanced sampling* dengan *cube method* dengan desain sampling yang telah ada, misalnya dengan *cluster sampling*, *multistage sampling*, ataupun *stratified sampling*.
2. Perlu dilakukan penelitian serupa di provinsi lain untuk menguji generalisasi dari hasil penelitian ini.
3. BPS perlu mempertimbangkan penerapan metode estimasi Bayesian karena hasil penelitian menunjukkan desain BPS dengan estimasi Bayesian menghasilkan estimator yang lebih valid, reliabel dan akurat. Penerapan metode Bayesian sangat memungkinkan karena BPS memiliki data yang sangat besar dan baik untuk dijadikan sebagai prior, sebagai contoh adalah data sensus, baik sensus penduduk, sensus pertanian, maupun sensus ekonomi. Selain itu, BPS juga memiliki data time series yang panjang untuk data yang rutin dikumpulkan.

DAFTAR PUSTAKA

Aggarwal, O.P. (1959), "Bayes and Minimax Procedures in Sampling from Finite and Infinite Population", *The Annals of Mathematical Statistics*, Vol. 30, hal. 206-218.

Aggarwal, O.P. (1966), "Bayes and Minimax Procedures for Estimating the Arithmetic Mean of a Population with Two Stage Sampling", *The Annals of Mathematical Statistics*, Vol. 37, hal. 1186-1195.

Allenby, G.M., Rossi, P.E. dan McCulloch, R.E. (2005), *Hierarchical Bayes Models: A Practitioners Guide*, viewed 05 February 2015, <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=655541>.

Ardilly, P. (1991), "Echantillonnage Representatif Optimum a Probabilites", *Annales d'Economie et de Statistique*, Vol. 23, hal. 91-113.

Basu, D. (1968), "The Role of The Sufficiency and The Likelihood Principles in Sample Survey Theory", *Sankhya: The Indian Journal of Statistics*, Vol. Series A, hal. 441-454.

Berger, J.O. (2005), *Introduction and History of Objective Bayesian Analysis, Tutorial on Objective Bayesian Analysis*, viewed 7 January 2015, <http://www.stat.missouri.edu/~bayes/Obayes5_webfiles/TutorialCourses/jim.pdf>.

Berger, Y.G. dan Tille, Y. (2009), "Sampling with Unequal Probabilities", dalam *Handbook of Statistics 29A*, Elsevier, Amsterdam.

Boldstat, W.M. (2010), *Understanding Computational Bayesian Statistics*, John Wiley & Sons, New Jersey.

BPS (2011), *Data Strategis BPS*, Jakarta.

BPS (2012), *Pedoman Pengawas Survei Angkatan Kerja Nasional (Sakernas)*, Jakarta.

Brewer, K. (2013), "Three Controversies in The History of Survey Sampling", *Survey Methodology*, hal. 249-262.

Brewer, K. dan Gregoire, T.G. (2009), "Introduction to Survey Sampling", dalam *Handbook of Statistics 29A*, eds. Pfeiffermann, D. dan Rao, C.R., Elsevier, Amsterdam.

- Carlin, B.P. dan Chib, S. (1995), "Bayesian Model Choice via Markov Chain Monte Carlo Methods", *Journal of the Royal Statistical Society*, Vol. 57, No. 3, hal. 473-484.
- Casella, G. dan George, E.I. (1992), "Explaining Gibbs Sampler", *The American Statistical Association*, Vol. 46, No. 3, hal. 167-174.
- Cochran, W.G. (1977), *Sampling Techniques Third Edition*, John Wiley & Sons, New York.
- Congdon, W.M. (2010), *Understanding Computational Bayesian Statistics*, John Wiley & Sons, New Jersey.
- Datta, G.S. (2009), "Model Based Approach to Small Area Estimation", dalam *Handbook of Statistics 29A*, eds. Pfeiffermann, D. dan Rao, C.R., Elsevier, Amsterdam.
- Deville, J.C. (1992), "Constrained Samples, Conditional Inference, Weighting: Three Aspects of The Utilisation of Auxiliary Information", *Proceedings of the Workshop on the Uses of Auxiliary Information in Surveys*, , Orebro (Sweden).
- Deville, J.C., Grosbras, J.M. dan Roth, N. (1988), "Efficient Sampling Algorithms and Balanced Sample", *COMPSTAT, Proceedings in Computational Statistics*, , Physica Verlag, Heidelberg, hal. 255-266.
- Deville, J.C. dan Tille, Y. (1998), "Unequal Probability Sampling Without Replacement Through a Splitting Method", *Biometrika*, Vol. 85, No. 1, hal. 89-101.
- Deville, J.C. dan Tille, Y. (2004), "Efficient Balanced Sampling: The Cube Method", *Biometrika*, hal. 893-912.
- Deville, J.C. dan Tille, Y. (2005), "Variance Approximation under Balanced Sampling", *Journal of Statistical Planning and Inference*, Vol. 128, hal. 569-591.
- Draper, N.R. dan Guttman, I. (1968), "Bayesian Stratified Two-Phase Sampling Result: k Characteristics", *Biometrika*, hal. 587-589.
- Ericson, W.A. (1965), "Optimum Stratified Sampling Using Prior Information", *Journal of American Statistical Association*, Vol. 60, hal. 750-771.

Ericson, W.A. (1969), "Subjective Bayesian Models in Sampling Finite Populations", *Journal of American Statistical Association*, Vol. 31, No. 2, hal. 195-233.

Geman, S. dan Geman, D. (1984), "Stochastic Relaxation, Gibbs Sampler, and The Bayesian Restoration of Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 6, No. 6, hal. 721-741.

Grosh, D.L. (1969), *Bayes One and Two Stage Stratified Sampling Schemes for Finite Populations with Beta-Binomial Prior Distributions*, Ph.D. Dissertation, Dept. of Statistics, Kansas State University.

Haziza, D. (2009), "Imputation Inference in the Presence of Missing Data", dalam *Handbook of Statistics 29A*, eds. Pfeiffermann, D. dan Rao, C.R., Elsevier, Amsterdam.

Hedayat, A.S. dan Majumdar, D. (1995), "Generating Desirable Sampling Plans by The Technique of Trade-off in Experimental Design", *Journal of Statistical Planning and Inference*, Vol. 44, hal. 237-247.

Iriawan, N. (2003), *Teknik Simulasi*, ITS, Surabaya.

Ismartini, P. (2013), *Pengembangan Model Linear Hirarki dengan Pendekatan Bayesian untuk Pemodelan Data Pengeluaran Perkapita Rumah Tangga*, Disertasi, ITS Surabaya, Surabaya.

Jordan, M.I. (2010), [www.cs.berkeley.edu](http://www.cs.berkeley.edu/~jordan/courses/260-spring10/lectures/lecture7.pdf), viewed 13 November 2014, <www.cs.berkeley.edu/~jordan/courses/260-spring10/lectures/lecture7.pdf>.

Kim, J.K., Brick, J.M., Fuller, W.A. dan Kalton, G. (2006), "On the bias of the multiple-imputation variance", *Journal of the Royal Statistical Society B*, Vol. 68, hal. 509-521.

Kott, P.S. (1995), "Aparadox of multiple imputation", *Proceedings of the Section on Survey Research Methods, American Statistical Association*, , hal. 380-383.

Krismawati dan Agustiyani, R. (2010), *Modul 8 Ketenagakerjaan Workshop Olah Cepat SP2010*, BPS, Jakarta.

Kusmayadi, E. (2011), *Perbandingan Teknik Penarikan Contoh untuk Menduga Hasil Pemilukada: Studi Kasus Pemilukada Kabupaten Jembrana*, Tesis, IPB Bogor, Bogor.

- Langel, M. dan Tille, L. (2011), "Corrado Gini, a Pioneer in Balanced Sampling and Inequality", *METRON – International Journal of Statistics*, Vol. LXIX, hal. 45-65.
- Legg, dan Yu, C.L.A. (2010), "Comparison of Sample Set Restriction Procedures", *Survey Methodology*, Vol. 36, No. 1, hal. 69-79.
- Levy, P.S. dan Lemeshow, S. (1999), *Sampling of Population: Methods and Application*, Third Edition Edition, Wiley Latherscience Publication, New York.
- Little, R.J.A. dan Rubin, D. (2002), *Statistical Analysis with Missing Data*, 2nd Edition, Wiley, New York.
- Lohr, S.L. (2009), "Part 1: Sampling and Survey Design, ", dalam *Handbook of Statistics 29A*, eds. Pfeffermann, D. dan Rao, C.R., Elsevier, Amsterdam.
- Meeden, G.A. (2012), "Bayesian Justification for Random Sampling in Sample Survey", *PJSOR*, Vol. 8, No. 3, hal. 353-357.
- Murphy, P. (2008), "An Overview of Primary Sampling Units (PSUs) in Multi-Stage Samples for Demographic Surveys", *Section on Government Statistics – Joint Statistical Meetings (JSM) 2008*, , Denver, hal. 2856-2863.
- Nedyalkova, D. dan Tille, Y. (2008), "Optimal Sampling and Estimation Strategies under the Linear Model", *Biometrika*, Vol. 95, No. 3, hal. 521-537.
- Neraca (2014), *Pengusaha Ragukan Data BPS*, viewed 15 October 2014, <<http://m.neraca.co.id/bisnis-indonesia/45095/Pengusaha-Ragukan-Data-BPS/>>.
- Neyman, J. (1934), "On The Different Aspects of The Representative Method: The Method of Stratified Sampling and The Method of Purposive Selection", *Journal of The Royal Statistical Society*, Vol. 97, No. 4, hal. 558-625.
- Ntzoufras, I. (2009), *Bayesian Modelling Using WinBUGS*, John Wiley & Sons, New Jersey.
- Opsomer, J. (2011), "Innovation in Survey Sampling Design: Discussion of Three Contributions Presented at U.S. Census Bureau, Statistics Canada", *Survey Methodology*, Vol. 37, No. 2, hal. 227-231.

Raghunathan, T.E. dan Rubin, D.B. (1997), "Role For Bayesian Technique in Survey Sampling", *SSC Annual Meeting, Proceedings of the Survey Methods Section*, , hal. 51-55.

Rao, J.N.K. (2011), "Impact of Frequentist and Bayesian Method on Survey Sampling Practice: A Selective Appraisal", *Statistical Science*, Vol. 26, No. 2, hal. 240-256.

Royall, R.M. dan Herson, J. (1973), "Robust Estimation in Finite Populations I", *Journal of the American Statistical Association*, Vol. 68, No. 344, hal. 880-889.

Rubin, D.B. (1987), *Multiple Imputations for Nonresponse in Surveys*, Wiley, New York.

Schafer, J.L. (1997), *Analysis of Incomplete Multivariate Data*, CRC Press, New York.

Thionet, P. (1953), *La Theorie des Sondages*, INSEE, Imprimerie nationale, Paris.

Tille, Y. (2006), *Sampling Algorithms*, Springer Science+Business Media, Inc, New York.

Tille, Y. (2011), "Ten Years of Balanced Sampling with Cube Method: An Appraisal", *Survey Methodology*, Vol. 37, No. 2, hal. 215-226.

Tille, Y. dan Favre, A.-C. (2004), "Coordination, Combination and Extension of Balanced Samples", *Biometrika*, Vol. 91, No. 4, hal. 913-927.

Tille, Y. dan Favre, A.-C. (2005), "Optimal Allocation in Balanced Sampling", *Statistics & Probability Letters*, Vol. 74, hal. 31–37.

Valliant, R., Dorfman, A.H. dan Royall, R.M. (2000), *Finite Population Sampling and Inference: A Prediction Approach*, Wiley, New York.

Widaningsih, R. (2008), "Perbandingan Berbagai Teknik Penarikan Contoh untuk Menduga Populasi Sapi Potong: Studi Kasus Kabupaten Karangasem, Propinsi Bali", *Informatika Pertanian*, Vol. 17, No. 2, hal. 1211-1230.

Yates, F. (1946), "A Review of Recent Statistical Developments in Sampling and Sampling Surveys", *Journal of the Royal Statistical Society*, No. A109, hal. 12-43.

Yates, F. (1949), *Sampling Methods for Censuses and Surveys*, Charles Griffin & Company Limited, London.

Zack, S. (1967), *Bayesian Design of Single, Double and Sequential Stratified Sampling for Estimating Proportion in Finite Population*, Tech. Report No. 27, Dept. of Statistics, Stanford University.

Zack, S. (1969), "Bayes Sequential Design of Fixed Size Samples from Finite Populations", *Journal of the American Statistical Association (JSTOR)*, Vol. 64, No. 328, hal. 1342-1349.

Zack, S. (1970), "Bayesian Design of Single and Double Stratified Sampling for Estimating Proportion in Finite Population", *Technometrics*, Vol. 12, No. 1, hal. 119-130.

LAMPIRAN

Lampiran 1. Pertanyaan Ketenagakerjaan pada Kuesioner Sensus Penduduk 2010

216. Kegiatan (**NAMA**) seminggu yang lalu:

a. Apakah bekerja atau berusaha?

Bekerja atau berusaha ialah melakukan kegiatan untuk memperoleh (membantu memperoleh) upah atau gaji atau laba yang dilakukan paling sedikit 1 jam.

☐ 1. Ya → ke P217

☐ 2. Tidak

b. Apakah mempunyai pekerjaan tetap tetapi sementara tidak bekerja?

Seperti menunggu panen, cuti, sakit, dll.

☐ 1. Ya → ke P217

☐ 2. Tidak

c. Apakah mencari pekerjaan atau mempersiapkan suatu usaha?

☐ 1. Ya → ke P219

☐ 2. Tidak

d. Apakah bersedia bekerja apabila ada yang menyediakan?

☐ 1. Ya

☐ 2. Tidak } ke P219

217. Apa lapangan usaha atau bidang pekerjaan (utama) dari tempat bekerja (**NAMA**) selama seminggu yang lalu?

(Tuliskan selengkap-lengkapnyanya, contoh: pertanian tanaman padi, sopir di perusahaan tekstil, sopir di Pemda, guru SMP Negeri, mengojek motor, dokter di Puskesmas, dsb).

.....
.....

Diisi oleh Kortim di lapangan

☐ 01. Pertanian tanaman padi & palawija (jagung, singkong, dll)

☐ 02. Hortikultura (sayur, buah, tanaman hias, tanaman obat, dll)

☐ 03. Perkebunan (tebu, teh, tembakau, karet, sawit, coklat, dll)

☐ 04. Perikanan (penangkapan, budidaya, biota laut, dll)

☐ 05. Peternakan (pembibitan & budidaya ternak besar/kecil, dll)

☐ 06. Kehutanan & pertanian lainnya (perburuan, sagu, rotan, dll)

☐ 07. Pertambangan & penggalian (pasir, emas, batubara, dll)

☐ 08. Industri pengolahan (anyaman, sepatu, pakaian, dll)

☐ 09. Listrik & gas (PLN, Non PLN, BM Gas, strom aki, dll)

Lampiran 2. Blok Sensus Sampel Sakernas Triwulanan 2011 Kalimantan Tengah
dengan Data Variabel Penelitian dari SP2010

No	kode_bs	Desa/ Kelurahan	X1	X2	X3	X4	X5	X6	X7	kota desa	Jml ruta
1	6201040013005	SUKAJAYA	237	0	237	167	17	39	14	2	122
2	6201050002007	KUMPAI BATU BAWAH	114	19	95	37	27	27	4	2	73
3	6201050004004	PASIR PANJANG	115	1	114	23	35	39	17	1	81
4	6201050007011	RAJA	201	14	187	16	14	97	60	1	101
5	6201050008016	SIDOREJO	295	26	269	17	40	103	109	1	152
6	6201050012004	RANGDA	289	108	181	28	106	27	20	2	122
7	6201050017011	MEDANGSARI	60	1	59	54	0	4	1	2	43
8	6201060013001	CANDI	242	13	229	29	33	71	96	1	138
9	6201070003004	SUKARAMI	171	0	171	167	2	2	0	2	74
10	6201070005002	KERABU	124	24	100	33	43	6	18	2	50
11	6202050004005	BAPINANG HULU	126	35	91	72	7	8	4	2	71
12	6202050010007	BABAUNG	229	54	175	130	11	17	17	2	97
13	6202060004066	MENTAWA BARU HILIR	191	0	191	39	51	48	53	1	119
14	6202060005004	MENTAWA BARU HULU	400	108	292	28	64	116	84	1	186
15	6202060005034		215	22	193	15	32	41	105	1	127
16	6202070002007	NATAI BARU	132	47	85	81	0	3	1	2	63
17	6202111004006	KENYALA	190	3	187	168	4	5	10	2	82
18	6202120002009	BAAMANG TENGAH	177	8	169	14	32	56	67	1	144
19	6202200003007	PARENGGEAN	194	25	169	24	12	89	44	1	92
20	6202200021001	BAJARAU	92	3	89	33	23	17	16	2	71
21	6202211004005	TUMBANG PENYAHUAN	234	29	205	172	6	17	10	2	111
22	6202230011002	TUMBANG BAJANEI	137	0	137	121	2	4	10	2	56
23	6203020003002	PALAMPAI	172	0	172	168	0	4	0	2	83
24	6203020006001	TAMBAN BARU SELATAN	158	7	151	145	0	3	3	2	73
25	6203030004001	ANJIR SERAPAT BARU	125	18	107	85	7	9	6	2	53
26	6203040013036	SELAT HULU	120	16	104	23	7	29	45	1	73
27	6203040015012	PULAU TELO	98	10	88	73	5	9	1	1	41
28	6203041007007	TERUSAN RAYA	204	0	204	198	1	4	1	2	77
29	6203070005003	LUNUK RAMBA	169	4	165	121	3	16	25	2	69
30	6203080002005	HAMPATUNG	142	0	142	67	6	15	54	1	88
31	6203090008002	ANJIR PALAMBANG	153	1	152	136	4	9	3	2	75
32	6203110002008	SAKA MANGKAHAI	117	9	108	77	5	5	21	2	66
33	6203150003002	UPT LAMUNTI C 1	83	0	83	63	7	5	8	2	35

No	kode_bs	Desa/ Kelurahan	X1	X2	X3	X4	X5	X6	X7	kota desa	Jml ruta
34	6204010007005	RANTAU KUJANG	106	14	92	67	0	10	15	2	81
35	6204020001015	SUNGAI JAYA	133	30	103	86	6	6	5	2	64
36	6204030002015	BANGKUANG	103	0	103	78	5	12	8	2	82
37	6204040006039	HILIR SPER	249	21	228	8	32	36	152	1	142
38	6204040014004	BUNTOK KOTA	143	11	132	1	14	97	20	1	59
39	6204040014035		139	7	132	1	40	18	73	1	76
40	6204040026005	TANJUNG JAWA	122	20	102	90	3	4	5	2	59
41	6204050018002	GUNUNG RANTAU	115	0	115	96	5	9	5	2	55
42	6205010002005	TUMPUNG LAUNG II	202	13	189	140	4	14	31	2	118
43	6205050001010	LEMO I	222	41	181	127	37	7	10	2	87
44	6205050013014	MELAYU	170	39	131	10	10	31	80	1	80
45	6205050013019		264	52	212	6	16	110	80	1	165
46	6205050013044		71	6	65	4	19	17	25	1	49
47	6205050016020	BUKIT SAWIT	171	0	171	166	0	5	0	2	79
48	6205050018002	SEI RAHAYU I	172	2	170	133	12	19	6	2	70
49	6205060001006	KARAMUAN	111	0	111	111	0	0	0	2	47
50	6206010006003	PULAU NIBUNG	113	1	112	67	17	17	11	2	56
51	6206020001012	MENDAWAI	232	50	182	28	50	42	62	1	134
52	6206020004004	PADANG	104	6	98	15	21	34	28	1	49
53	6206020005004	KARTA MULIA	121	9	112	74	18	12	8	2	64
54	6206020008002	PETARIKAN	210	3	207	156	2	36	13	2	111
55	6207010005013	NANGA BULIK	227	16	211	13	14	45	139	1	126
56	6207010005028		170	12	158	98	4	39	17	1	92
57	6207010006003	BUMI AGUNG	190	64	126	29	21	35	41	2	102
58	6207012007005	MODANG MAS	207	59	148	112	7	19	10	2	92
59	6207013001005	BUKIT JAYA	138	10	128	50	4	45	29	2	83
60	6207020004003	SUJA	249	19	230	206	3	9	12	2	118
61	6207020006005	BAKONSU	219	65	154	115	1	17	21	2	81
62	6208010002014	KUALA PEMBUANG DUA	152	8	144	9	14	31	90	1	83
63	6208010005003	KUALA PEMBUANG SATU	139	33	106	15	24	18	49	1	62
64	6208010010001	BAUNG	226	50	176	155	1	12	8	2	105
65	6208020007002	BANUA USANG	118	0	118	110	0	6	2	2	65
66	6208030004005	PEMBUANG HULU I	265	24	241	123	45	49	24	1	129
67	6208040012021	BATU AGUNG	92	8	84	61	9	9	5	2	46
68	6208040020007	SUKA MANDANG	206	43	163	117	14	17	15	2	74
69	6208050025001	TUMBANG LAKU	94	50	44	43	1	0	0	2	41
70	6209010015004	BUMI SUBUR	152	2	150	143	0	2	5	2	91

No	kode_bs	Desa/ Kelurahan	X1	X2	X3	X4	X5	X6	X7	kota desa	Jml ruta
71	6209050002004	TUMBANG LITING	148	25	123	66	12	6	39	2	63
72	6209050003018	KASONGAN BARU	114	2	112	10	10	15	77	1	63
73	6209050008014	HAMPALIT	155	29	126	14	37	40	35	1	80
74	6209050008024		148	17	131	2	29	85	15	1	77
75	6209060004004	BANGKUANG	108	12	96	51	39	3	3	2	53
76	6209060009006	TEWANG RANGKANG	202	34	168	100	41	19	8	2	100
77	6209090010005	TUMBANG ATEI	136	27	109	56	27	9	17	2	58
78	6209111008009	TUMBANG	98	18	80	75	1	2	2	2	80
79	6210010006005	SEI RUNGUN	110	8	102	82	6	4	10	2	62
80	6210020007010	TALIO HULU	137	6	131	79	28	8	16	2	81
81	6210030014010	GANDANG BARAT	104	0	104	72	19	9	4	2	55
82	6210040005008	PULANG PISAU	194	3	191	18	30	45	98	1	117
83	6210040005014		170	8	162	32	26	22	82	1	87
84	6210050007001	BUKIT LITI	286	4	282	213	39	12	18	2	123
85	6210050009001	PAMARUNAN	113	4	109	105	2	1	1	2	55
86	6210060005001	HANUA	220	0	220	128	24	14	54	2	87
87	6211010010005	TANGKI DAHUYAN	228	4	224	224	0	0	0	2	125
88	6211020009004	LINAU	164	0	164	156	0	5	3	2	69
89	6211020013008	RABAMBANG	174	1	173	148	2	12	11	2	73
90	6211021007004	TUMBANG RAHUYAN	107	6	101	61	19	5	16	2	49
91	6211030004005	SEPANG SIMIN	289	17	272	193	18	38	23	1	141
92	6211050004014	TEWAH	121	14	107	11	28	34	34	1	58
93	6211062003004	TUMBANG NAPOI	155	7	148	126	11	2	9	2	58
94	6212020006001	TAMIANG LAYANG	129	18	111	10	30	29	42	1	65
95	6212050005026	RODOK	318	9	309	233	17	42	17	1	125
96	6212050005029		188	0	188	164	2	13	9	1	95
97	6212060001002	MUARA PALANTAU	149	1	148	120	2	17	9	2	104
98	6212070007002	TAMPU LANGIT	276	31	245	194	15	18	18	2	94
99	6212090003003	PATUNG	98	4	94	56	18	6	14	2	45
100	6212090007001	LUAU JAWUK	121	41	80	63	12	3	2	2	52
101	6213010005003	MUARA BAKANON	293	33	260	174	39	40	7	2	158
102	6213020008004	BAHITOM	189	0	189	177	0	12	0	2	74
103	6213020009002	DANAU USUNG	167	9	158	106	1	27	24	2	92
104	6213020011001	BERIWIT	173	4	169	0	13	86	70	1	87
105	6213020011025		668	8	660	22	88	100	450	1	376
106	6213020012001	PURUK CAHU	261	0	261	175	4	65	17	2	116

No	kode_bs	Desa/ Kelurahan	X1	X2	X3	X4	X5	X6	X7	kota desa	Jml ruta
107	6213050006002	TELUK JOLO	154	54	100	43	35	8	14	2	57
108	6271010004020	PAHANDUT	181	105	76	0	13	45	18	1	88
109	6271011003003	KALAMPANGAN	295	30	265	163	38	33	31	1	118
110	6271012001057	MENTENG	246	35	211	11	32	35	133	1	195
111	6271012002040	PALANGKA	73	10	63	0	14	19	30	1	86
112	6271012002063		227	40	187	10	23	81	73	1	111
113	6271012002123		131	3	128	4	43	54	27	1	86
114	6271012003042	BUKIT TUNGGAL	186	8	178	4	25	28	121	1	112
115	6271021005002	PETUK BERUNAI	191	8	183	31	59	72	21	2	103
116	6271021006002	MUNGKU BARU	141	8	133	66	45	12	10	2	57

Lampiran 3. Syntax Program R Penarikan Sampel *Balanced Sampling* dengan *Cube Method*

```
# matrix of balancing variables
X=cbind(kalteng$X1,
        kalteng$X2,
        kalteng$X3,
        kalteng$X4,
        kalteng$X5,
        kalteng$X6,
        kalteng$X7)
# vector of inclusion probabilities n sample size is 116
pik=inclusionprobabilities(kalteng$jml_pddk,116)
# selection of the sample
sample=samplecube(X,pik,order=1,comment=TRUE)
# The selected sample
options(digits=13)
as.numeric(kalteng$Kode_BS[sample==1])
as.numeric(kalteng$X1[sample==1])
as.numeric(kalteng$X2[sample==1])
as.numeric(kalteng$X3[sample==1])
as.numeric(kalteng$X4[sample==1])
as.numeric(kalteng$X5[sample==1])
as.numeric(kalteng$X6[sample==1])
as.numeric(kalteng$X7[sample==1])
```


Lampiran 4. Hasil Penarikan Sampel dengan Pendekatan *Balanced Sampling* dengan *Cube Method*

```
> # matrix of balancing variables
> X=cbind(kal teng$X1,
+         kal teng$X2,
+         kal teng$X3,
+         kal teng$X4,
+         kal teng$X5,
+         kal teng$X6,
+         kal teng$X7)
> # vector of inclusion probabilities n sample size is 116
> pi k=inclusionprobabilities(kal teng$jml_pddk, 116)
> # selection of the sample
> sample=sampecube(X, pi k, order=1, comment=TRUE)
```

BEGINNING OF THE FLIGHT PHASE

The matrix of balanced variable has 7 variables and 7719 units
 The size of the inclusion probability vector is 7719
 The sum of the inclusion probability vector is 116
 The inclusion probability vector has 7719 non-integer elements
 Step 1 Step 2 ,

BEGINNING OF THE LANDING PHASE

At the end of the flight phase, there remain 5 non integer probabilities
 The sum of these probabilities is 2.89015924526
 This sum is non-integer
 The linear program will consider 20 possible samples
 The mean cost is 0.003307542213057
 The smallest cost is 0.0009751994611028
 The largest cost is 0.007599887650772
 The cost of the selected sample is 0.0009751994611028

QUALITY OF BALANCING

No	TOTALS	HorvitzThompson_estimators	Relative_deviation
1	1145833	1148324.1330528	0.2174080387653
2	101350	101473.5612993	0.1219154408313
3	1044483	1046850.5717535	0.2266740342871
4	597862	601015.3405796	0.5274361942358
5	118288	117985.8719256	-0.2554173495439
6	157868	159880.8392033	1.2750140644993
7	170465	167968.5200450	-1.4645117501855

```
> # The selected sample
> options(digits=13)
> as.numeric(kal teng$Kode_BS[sample==1])
[1] 6201040008001 6201040010002 6201040010004 6201050004003 620
1050007006 6201050010001
[7] 6201060009008 6201060014007 6201060014013 6201060031008 620
1061007005 6201061009001
[13] 6201061013001 6201062006009 6201062010003 6201070003002 620
2060004058 6202060004062
[19] 6202060005042 6202060006001 6202061004001 6202070003008 620
2070005002 6202070006002
[25] 6202070006003 6202120002027 6202120002077 6202120003015 620
2190007001 6202190007003
[31] 6202190008001 6202191003003 6202191008003 6202200010004 620
2200018003 6202200021003
[37] 6202210003009 6202210010011 6202210012022 6202230013005 620
3020006006 6203030003008
[43] 6203040012035 6203040014010 6203070003003 6203070004003 620
3090003006 6203090004010
```



```

[49] 6203100006006 6203100008001 6203100033003 6203150024001 620
3160006006 6204020001010
[55] 6204030002013 6204030002014 6204030005005 6204040003012 620
4040005003 6204060015001
[61] 6205010010002 6205050005010 6205050013042 6205060004004 620
6020004008 6206020005010
[67] 6206030006003 6206031002002 6207013003004 6208010005002 620
8010016003 6208010016004
[73] 6208010016005 6208010016006 6208020001002 6208030006005 620
8030013007 6208040007003
[79] 6208050010005 6209010002006 6209010011002 6209040005002 620
9050003014 6209050008018
[85] 6209050008028 6209050008034 6210010010003 6210020004020 621
0030007013 6210040001004
[91] 6210040006012 6210040007004 6210041001002 6211020002003 621
1021004004 6211040011004
[97] 6211050004016 6211061004002 6211061008004 6212030004002 621
2060004003 6213030030005
[103] 6213040010003 6213040025001 6213040032001 6213041004003 621
3050008002 6271010005013
[109] 6271010006054 6271012001001 6271012001067 6271012002037 627
1012002049 6271012002113
[115] 6271012003102 6271020003005
> as.numeric(kal teng$X1[sample==1])
[1] 221 225 194 252 284 171 136 167 231 139 148 273 203 123 202
142 143 134 310 290 107 129
[23] 114 201 124 144 163 138 160 166 133 171 188 162 115 132 450
85 353 206 61 192 176 165
[45] 104 151 188 109 197 290 91 134 154 77 151 116 84 164 204
143 88 183 86 377 138 102
[67] 134 143 91 237 138 86 171 236 128 195 200 250 140 237 152
100 47 250 207 184 201 136
[89] 186 286 91 192 145 124 232 274 209 185 182 69 150 180 163
238 360 159 188 165 89 290
[111] 246 54 257 181 257 149
> as.numeric(kal teng$X2[sample==1])
[1] 18 21 20 42 21 27 8 4 70 1 2 0 7 0 11 4 7 8 3 66
9 34 33 2 0 5 10 3 5
[30] 2 0 3 1 2 0 7 11 7 0 12 0 20 15 18 1 3 17 6 0
0 0 8 40 9 18 10 16 12
[59] 22 46 7 0 17 3 22 1 10 3 8 46 1 0 4 7 49 20 5 11
30 30 15 17 12 36 1 54 0
[88] 1 6 0 6 9 38 15 1 26 65 22 2 0 7 4 0 44 15 12 32
22 26 4 57 3 14 20 66 28
> as.numeric(kal teng$X3[sample==1])
[1] 203 204 174 210 263 144 128 163 161 138 146 273 196 123 191
138 136 126 307 224 98 95
[23] 81 199 124 139 153 135 155 164 133 168 187 160 115 125 439
78 353 194 61 172 161 147
[45] 103 148 171 103 197 290 91 126 114 68 133 106 68 152 182
97 81 183 69 374 116 101
[67] 124 140 83 191 137 86 167 229 79 175 195 239 110 207 137
83 35 214 206 130 201 135
[89] 180 286 85 183 107 109 231 248 144 163 180 69 143 176 163
194 345 147 156 143 63 286
[111] 189 51 243 161 191 121
> as.numeric(kal teng$X4[sample==1])
[1] 165 140 113 4 20 16 40 3 13 33 143 244 146 63 113
130 13 7 12 46 78 69
[23] 37 115 75 14 22 34 145 146 122 107 163 120 84 51 376
49 342 171 60 109 14 4
[45] 39 109 142 57 192 240 91 83 26 55 105 59 31 73 120
70 57 161 0 358 28 85
[67] 51 106 68 5 78 71 105 94 75 111 195 223 73 161 132
52 11 18 3 0 154 133

```



```

[89] 76 237 43 150 52 86 194 47 46 99 46 57 103 119 137
175 252 92 90 1 0 16
[111] 6 1 5 4 37 6
> as.numeric(kal teng$X5[sample==1])
[1] 9 16 10 10 17 11 22 12 79 27 0 7 4 1 44
0 15 36 65 73 2 3
[23] 3 30 16 22 16 42 0 2 1 19 3 9 3 7 63
18 1 11 0 10 21 9
[45] 20 5 1 12 1 42 0 1 75 0 6 17 26 54 25
12 6 0 5 0 35 1
[67] 4 1 1 46 3 7 8 17 0 30 0 7 26 15 0
24 3 77 11 74 6 0
[89] 15 11 2 17 5 15 21 81 64 24 118 1 3 21 1
4 32 21 39 10 10 48
[111] 34 11 25 43 6 56
> as.numeric(kal teng$X6[sample==1])
[1] 8 38 31 61 142 70 32 124 42 50 0 11 34 23 19
1 71 49 138 36 11 8
[23] 11 28 19 52 77 42 6 9 7 10 11 17 18 50 0
5 1 7 1 25 93 23
[45] 25 27 16 21 1 5 0 14 12 6 13 16 6 17 26
5 6 17 25 2 31 1
[67] 19 10 2 46 7 4 18 20 3 16 0 9 7 13 1
3 7 78 170 47 12 0
[89] 63 17 24 10 23 4 10 33 20 8 11 3 8 17 14
3 18 7 9 51 33 51
[111] 63 9 40 75 32 31
> as.numeric(kal teng$X7[sample==1])
[1] 21 10 20 135 84 47 34 24 27 28 3 11 12 36 15
7 37 34 92 69 7 15
[23] 30 26 14 51 38 17 4 7 3 32 10 14 10 17 0
6 9 5 0 28 33 111
[45] 19 7 12 13 3 3 0 28 1 7 9 14 5 8 11
10 12 5 39 14 22 14
[67] 50 23 12 94 49 4 36 98 1 18 0 0 4 18 4
4 14 41 22 9 29 2
[89] 26 21 16 6 27 4 6 87 14 32 5 8 29 19 11
12 43 27 18 81 20 171
[111] 86 30 173 39 116 28

```


Lampiran 5. Blok Sensus dan Data Variabel Penelitian dari Sampel *Balanced Sampling* dengan *Cube Method*

No	kode_bs	Desa/ Kelurahan	X1	X2	X3	X4	X5	X6	X7
1	6201040008001	KOTAWARINGIN HULU	221	18	203	165	9	8	21
2	6201040010002	RUNGUN	225	21	204	140	16	38	10
3	6201040010004		194	20	174	113	10	31	20
4	6201050004003	PASIR PANJANG	252	42	210	4	10	61	135
5	6201050007006	RAJA	284	21	263	20	17	142	84
6	6201050010001	BARU	171	27	144	16	11	70	47
7	6201060009008	SUNGAI KAPITAN	136	8	128	40	22	32	34
8	6201060014007	KUMAI HULU	167	4	163	3	12	124	24
9	6201060014013		231	70	161	13	79	42	27
10	6201060031008	BUMI HARJO	139	1	138	33	27	50	28
11	6201061007005	ARGA MULYA	148	2	146	143	0	0	3
12	6201061009001	SIDO MULYO	273	0	273	244	7	11	11
13	6201061013001	SUNGAI KUNING	203	7	196	146	4	34	12
14	6201062006009	PANDU SANJAYA	123	0	123	63	1	23	36
15	6201062010003	PANGKALAN DURIN	202	11	191	113	44	19	15
16	6201070003002	SUKARAMI	142	4	138	130	0	1	7
17	6202060004058	MENTAWA BARU HILIR	143	7	136	13	15	71	37
18	6202060004062		134	8	126	7	36	49	34
19	6202060005042	MENTAWA BARU HULU	310	3	307	12	65	138	92
20	6202060006001	SAWAHAN	290	66	224	46	73	36	69
21	6202061004001	TERANTANG	107	9	98	78	2	11	7
22	6202070003008	BAGENDANG TENGAH	129	34	95	69	3	8	15
23	6202070005002	BAGENDANG HULU	114	33	81	37	3	11	30
24	6202070006002	SUMBER MAKMUR	201	2	199	115	30	28	26
25	6202070006003		124	0	124	75	16	19	14
26	6202120002027	BAAMANG TENGAH	144	5	139	14	22	52	51
27	6202120002077		163	10	153	22	16	77	38
28	6202120003015	BAAMANG HULU	138	3	135	34	42	42	17
29	6202190007001	PATAI	160	5	155	145	0	6	4
30	6202190007003		166	2	164	146	2	9	7
31	6202190008001	RUBUNG BUYUNG	133	0	133	122	1	7	3
32	6202191003003	KERUING	171	3	168	107	19	10	32

No	kode_bs	Desa/ Kelurahan	X1	X2	X3	X4	X5	X6	X7
33	6202191008003	BUKIT BATU	188	1	187	163	3	11	10
34	6202200010004	BARINGIN TUNGGAL JAYA	162	2	160	120	9	17	14
35	6202200018003	DAMAR MAKMUR	115	0	115	84	3	18	10
36	6202200021003	BAJARAU	132	7	125	51	7	50	17
37	6202210003009	KAWAN BATU	450	11	439	376	63	0	0
38	6202210010011	PEMANTANG	85	7	78	49	18	5	6
39	6202210012022	KUALA KUAYAN	353	0	353	342	1	1	9
40	6202230013005	TUMBANG SEPAYANG	206	12	194	171	11	7	5
41	6203020006006	TAMBAK BARU SELATAN	61	0	61	60	0	1	0
42	6203030003008	ANJIR SERAPAT BARAT	192	20	172	109	10	25	28
43	6203040012035	SELAT TENGAH	176	15	161	14	21	93	33
44	6203040014010	SELAT DALAM	165	18	147	4	9	23	111
45	6203070003003	MALUEN	104	1	103	39	20	25	19
46	6203070004003	BASUNGKAI	151	3	148	109	5	27	7
47	6203090003006	NARAHAN	188	17	171	142	1	16	12
48	6203090004010	BUNGA MAWAR	109	6	103	57	12	21	13
49	6203100006006	TAJEPAN	197	0	197	192	1	1	3
50	6203100008001	MUARA DADAHUP	290	0	290	240	42	5	3
51	6203100033003	UPT DADAHUP C2	91	0	91	91	0	0	0
52	6203150024001	KALUMPANG	134	8	126	83	1	14	28
53	6203160006006	DANAU PANTAU	154	40	114	26	75	12	1
54	6204020001010	SUNGAI JAYA	77	9	68	55	0	6	7
55	6204030002013	BANGKUANG	151	18	133	105	6	13	9
56	6204030002014		116	10	106	59	17	16	14
57	6204030005005	BABAI	84	16	68	31	26	6	5
58	6204040003012	BARU	164	12	152	73	54	17	8
59	6204040005003	JELAPAT	204	22	182	120	25	26	11
60	6204060015001	PATAS II	143	46	97	70	12	5	10
61	6205010010002	KAMAWEN	88	7	81	57	6	6	12
62	6205050005010	SABUH	183	0	183	161	0	17	5
63	6205050013042	MELAYU	86	17	69	0	5	25	39
64	6205060004004	TELUK MALEWAI	377	3	374	358	0	2	14
65	6206020004008	PADANG	138	22	116	28	35	31	22
66	6206020005010	KARTA MULIA	102	1	101	85	1	1	14
67	6206030006003	BALAI RIAM	134	10	124	51	4	19	50
68	6206031002002	AJANG	143	3	140	106	1	10	23
69	6207013003004	NANGA PALIKODAN	91	8	83	68	1	2	12

No	kode_bs	Desa/ Kelurahan	X1	X2	X3	X4	X5	X6	X7
70	6208010005002	KUALA PEMBUANG	237	46	191	5	46	46	94
71	6208010016003	SUNGAI UNDANG	138	1	137	78	3	7	49
72	6208010016004		86	0	86	71	7	4	4
73	6208010016005		171	4	167	105	8	18	36
74	6208010016006		236	7	229	94	17	20	98
75	6208020001002	TELAGA PULANG	128	49	79	75	0	3	1
76	6208030006005	DERANGGA	195	20	175	111	30	16	18
77	6208030013007	RUNGAU RAYA	200	5	195	195	0	0	0
78	6208040007003	KALANG	250	11	239	223	7	9	0
79	6208050010005	TUMBANG MANJUL	140	30	110	73	26	7	4
80	6209010002006	PEGATAN HULU	237	30	207	161	15	13	18
81	6209010011002	SEBANGUN JAYA	152	15	137	132	0	1	4
82	6209040005002	TUMBANG PANGGO	100	17	83	52	24	3	4
83	6209050003014	KASONGAN BARU	47	12	35	11	3	7	14
84	6209050008018	HAMPALIT	250	36	214	18	77	78	41
85	6209050008028		207	1	206	3	11	170	22
86	6209050008034		184	54	130	0	74	47	9
87	6210010010003	TANJUNG PERAWAN	201	0	201	154	6	12	29
88	6210020004020	BELANTI SIAM	136	1	135	133	0	0	2
89	6210030007013	KANAMIT	186	6	180	76	15	63	26
90	6210040001004	BUNTOI	286	0	286	237	11	17	21
91	6210040006012	ANJIR PULANG PISAU	91	6	85	43	2	24	16
92	6210040007004	GOHONG	192	9	183	150	17	10	6
93	6210041001002	GARONG	145	38	107	52	5	23	27
94	6211020002003	TALANGKAH	124	15	109	86	15	4	4
95	6211021004004	SANGAL	232	1	231	194	21	10	6
96	6211040011004	TAMPANG TUMBANG	274	26	248	47	81	33	87
97	6211050004016	TEWAH	209	65	144	46	64	20	14
98	6211061004002	TUMBANG MARIKOI	185	22	163	99	24	8	32
99	6211061008004	TUMBANG MAHUROI	182	2	180	46	118	11	5
100	6212030004002	AMPARI	69	0	69	57	1	3	8
101	6212060004003	TUYAU	150	7	143	103	3	8	29
102	6213030030005	BATU BUA I	180	4	176	119	21	17	19
103	6213040010003	DIRUNG BAKUNG	163	0	163	137	1	14	11
104	6213040025001	MUWUN	238	44	194	175	4	3	12
105	6213040032001	KARALI	360	15	345	252	32	18	43

No	kode_bs	Desa/ Kelurahan	X1	X2	X3	X4	X5	X6	X7
106	6213041004003	DIRUNG LINGKIN	159	12	147	92	21	7	27
107	6213050008002	TUMBANG TUAN	188	32	156	90	39	9	18
108	6271010005013	PANARUNG	165	22	143	1	10	51	81
109	6271010006054	LANGKAI	89	26	63	0	10	33	20
110	6271012001001	MENTENG	290	4	286	16	48	51	171
111	6271012001067		246	57	189	6	34	63	86
112	6271012002037	PALANGKA	54	3	51	1	11	9	30
113	6271012002049		257	14	243	5	25	40	173
114	6271012002113		181	20	161	4	43	75	39
115	6271012003102	BUKIT TUNGGAL	257	66	191	37	6	32	116
116	6271020003005	BANTURUNG	149	28	121	6	56	31	28

Lampiran 6. Syntax Program R untuk Penarikan Sampel dengan Pendekatan

Reference Prior

```
sample_kab<-NULL
sample_kec<-NULL
sample_desa<-NULL
sample_bs<-NULL
kab$kurang_bobot<-c(0)
kec$kurang_bobot<-c(0)
desa$kurang_bobot<-c(0)
bs$kurang_bobot<-c(0)
d<-1
n<-116
sample_gab=NULL
for (d in 1:n){
  kabsample<-kab
  nkab<-length(kabsample$kode_kab)
  kum1=cumsum(kabsample$X2-kabsample$kurang_bobot)
  total_pengangguran_prov=kum1[nkab]
  u[d]=runif(1,0,total_pengangguran_prov)
  e<-1
  while (kum1[e] < u[d]) {
    e<-e+1}
  e
  sample_kab[d]=kabsample$kode_kab[e]
  kecsample<-subset(kec,kec$kode_kab == kab$kode_kab[e])
  nkec<-length(kecsample$kode_kab)
  kum2=cumsum(kecsample$X2-kecsample$kurang_bobot)
  total_pengangguran_kab=kum2[nkec]
  v[d]=runif(1,0,total_pengangguran_kab)
  f<-1
  while (kum2[f] < v[d]) {
    f<-f+1}
  f
  sample_kec[d]=kecsample$kode_kec[f]
  desasample<-subset(desasample,desasample$kode_kec == kecsample$kode_kec[f])
  ndesa<-length(desasample$kode_kec)
  kum3=cumsum(desasample$X2-desasample$kurang_bobot)
  total_pengangguran_kec=kum3[ndesa]
  j[d]=runif(1,0,total_pengangguran_kec)
  g<-1
  while (kum3[g] < j[d]) {
    g<-g+1}
  g
  sample_desa[d]=desasample$kode_desa[g]
  bssample<-subset(bs,bs$kode_desa == desasample$kode_desa[g])
  nbs<-length(bssample$kode_desa)
  kum4=cumsum(bssample$X2-bssample$kurang_bobot)
  total_pengangguran_desa=kum4[nbs]
  k[d]=runif(1,0,total_pengangguran_desa)
```



```

h<-1
while (kum4[h] < k[d]) {
  h<-h+1}
h
sample_bs[d]<-bssample$kode_bs[h]
kab$kurang_bobot[e]<-kab$kurang_bobot[e]+bssample$X2[h]
kec$kurang_bobot[which(kec$kode_kec==kecsample$kode_kec[f])<-
kec$kurang_bobot[which(kec$kode_kec==kecsample$kode_kec[f])+bssample$
X2[h]
desa$kurang_bobot[which(desa$kode_desa==desasample$kode_desa[g])<-
desa$kurang_bobot[which(desasample$kode_desa[g])+bssample$X2[h]
bs$kurang_bobot[which(bs$kode_bs==bssample$kode_bs[h])<-
bs$kurang_bobot[which(bs$kode_bs==bssample$kode_bs[h])+bssample$X2[h]
}
options(digits=13)
sample_gab=cbind(sample_kab, sample_kec, sample_desa, sample_bs)
sample_bs

```


Lampiran 7. Hasil Penarikan Sampel dengan Pendekatan *Reference Prior*

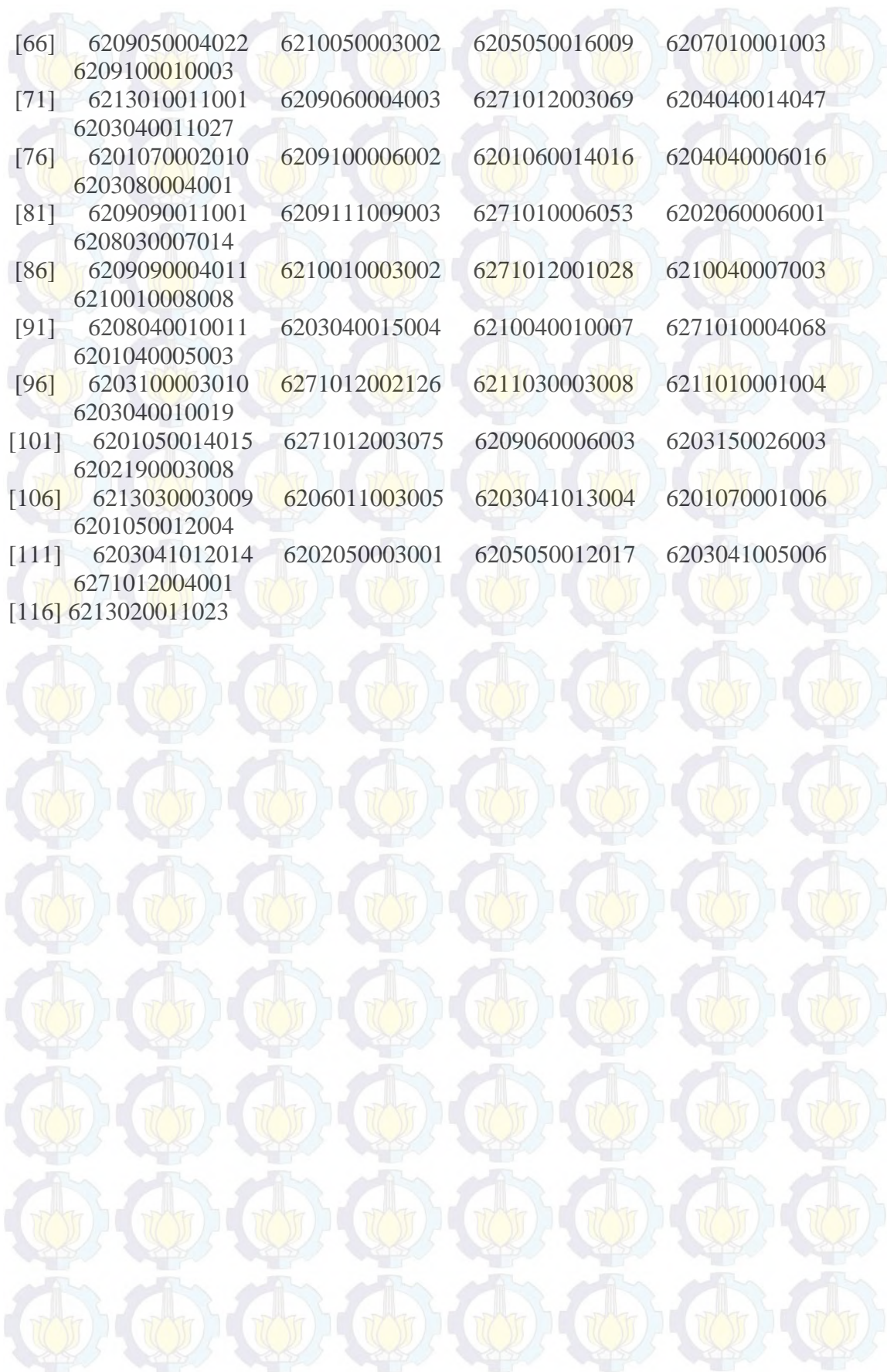
```
> sample_kab<-NULL
> sample_kec<-NULL
> sample_desa<-NULL
> sample_bs<-NULL
> kab$kurang_bobot<-c(0)
> kec$kurang_bobot<-c(0)
> desa$kurang_bobot<-c(0)
> bs$kurang_bobot<-c(0)
> d<-1
> n<-116
> sample_gab=NULL
> for (d in 1:n){
+   kabsample<-kab
+   nkab<-length(kabsample$kode_kab)
+   kum1=cumsum(kabsample$X2-kabsample$kurang_bobot)
+   total_pengangguran_prov=kum1[nkab]
+   u[d]=runif(1,0,total_pengangguran_prov)
+   e<-1
+   while (kum1[e] < u[d]) {
+     e<-e+1
+   }
+   e
+   sample_kab[d]=kabsample$kode_kab[e]
+   kecsample<-subset(kec,kec$kode_kab == kab$kode_kab[e])
+   nkec<-length(kecsample$kode_kab)
+   kum2=cumsum(kecsample$X2-kecsample$kurang_bobot)
+   total_pengangguran_kab=kum2[nkec]
+   v[d]=runif(1,0,total_pengangguran_kab)
+   f<-1
+   while (kum2[f] < v[d]) {
+     f<-f+1
+   }
+   f
+   sample_kec[d]=kecsample$kode_kec[f]
+   desasample<-subset(desa,desa$kode_kec == kecsample$kode_kec[f])
+   ndesa<-length(desasample$kode_kec)
+   kum3=cumsum(desasample$X2-desasample$kurang_bobot)
+   total_pengangguran_kec=kum3[ndesa]
+   j[d]=runif(1,0,total_pengangguran_kec)
+   g<-1
+   while (kum3[g] < j[d]) {
+     g<-g+1
+   }
+   g
+   sample_desa[d]=desasample$kode_desa[g]
+   bssample<-subset(bs,bs$kode_desa == desasample$kode_desa[g])
+   nbs<-length(bssample$kode_desa)
+   kum4=cumsum(bssample$X2-bssample$kurang_bobot)
```



```

+ total_pengangguran_desa=kum4[nbs]
+ k[d]=runif(1,0,total_pengangguran_desa)
+ h<-1
+ while (kum4[h] < k[d]) {
+   h<-h+1}
+ h
+ sample_bs[d]<-bssample$kode_bs[h]
+ kab$kurang_bobot[e]<-kab$kurang_bobot[e]+bssample$X2[h]
+   kec$kurang_bobot[which(kec$kode_kec==kecsample$kode_kec[f])<-
kec$kurang_bobot[which(kec$kode_kec==kecsample$kode_kec[f])+bssample$
X2[h]
+   desa$kurang_bobot[which(desa$kode_desa==desasample$kode_desa[g])<-
desa$kurang_bobot[which(desa$kode_desa==desasample$kode_desa[g])+bssam
ple$X2[h]
+   bs$kurang_bobot[which(bs$kode_bs==bssample$kode_bs[h])<-
bs$kurang_bobot[which(bs$kode_bs==bssample$kode_bs[h])+bssample$X2[h]
+ }
> options(digits=13)
> sample_gab=cbind(sample_kab, sample_kec, sample_desa, sample_bs)
> sample_bs
[1] 6202120003015 6208020012006 6202060007007 6202060003033
6209100013002
[6] 6210011002002 6202120002106 6205010002009 6202211004003
6209050004013
[11] 6271021001005 6271010004062 6207021006007 6208040026001
6201060010019
[16] 6207020010004 6201050010045 6204040004001 6204040026007
6271010006012
[21] 6271012003021 6202190003012 6208010005014 6271012003060
6203150019001
[26] 6204030002014 6213040032001 6271012001064 6202060003009
6211060008002
[31] 6202020008003 6210040010006 6271012001009 6213052003002
6203170006004
[36] 6207010005008 6209030003003 6209050008035 6271010011010
6202060005004
[41] 6206020004003 6210020015008 6204020006004 6203150019011
6203040013013
[46] 6202021002005 6209080001002 6203180018001 6201060006001
6271012002004
[51] 6203080001002 6202050007003 6271012002038 6201061005011
6211040012018
[56] 6201050005028 6210060011001 6208010005002 6203110005002
6210040003008
[61] 6209060010002 6212050009003 6212050007002 6204030004002
6211060011002

```

[66]	6209050004022 6209100010003	6210050003002	6205050016009	6207010001003
[71]	6213010011001 6203040011027	6209060004003	6271012003069	6204040014047
[76]	6201070002010 6203080004001	6209100006002	6201060014016	6204040006016
[81]	6209090011001 6208030007014	6209111009003	6271010006053	6202060006001
[86]	6209090004011 6210010008008	6210010003002	6271012001028	6210040007003
[91]	6208040010011 6201040005003	6203040015004	6210040010007	6271010004068
[96]	6203100003010 6203040010019	6271012002126	6211030003008	6211010001004
[101]	6201050014015 6202190003008	6271012003075	6209060006003	6203150026003
[106]	6213030003009 6201050012004	6206011003005	6203041013004	6201070001006
[111]	6203041012014 6271012004001	6202050003001	6205050012017	6203041005006
[116]	6213020011023			

Lampiran 8. Blok Sensus dan Data Variabel Penelitian dari Sampel *Reference Prior Sampling*

No	kode_bs	Desa/ Kelurahan	X1	X2	X3	X4	X5	X6	X7
1	6201040005003	KOTA WARINGIN HILIR	253	57	196	77	17	64	38
2	6201050005028	MENDAWAI	130	34	96	4	5	68	19
3	6201050010045	BARU	205	22	183	147	12	15	9
4	6201050012004	RANGDA	289	108	181	28	106	27	20
5	6201050014015	RUNTU	554	62	492	492	0	0	0
6	6201060006001	TELUK BOGAM	175	25	150	96	40	13	1
7	6201060010019	KUMAI HILIR	191	49	142	28	45	37	32
8	6201060014016	KUMAI HULU	173	15	158	25	80	38	15
9	6201061005011	KARANGMULYA	205	22	183	54	40	68	21
10	6201070001006	NANGA MUA	229	32	197	175	1	2	19
11	6201070002010	PANGKUT	252	2	250	248	0	2	0
12	6202020008003	SAMUDA KOTA	177	91	86	41	22	18	5
13	6202021002005	LAMPUYANG	175	2	173	160	0	0	13
14	6202050003001	BAPINANG HILIR	138	32	106	85	8	4	9
15	6202050007003	BABIRAH	213	105	108	71	21	7	9
16	6202060003009	KETAPANG	280	4	276	43	58	141	34
17	6202060003033		172	8	164	19	33	60	52
18	6202060005004	MENTAWA BARU HULU	400	108	292	28	64	116	84
19	6202060006001	SAWAHAN	290	66	224	46	73	36	69
20	6202060007007	BAPEANG	106	36	70	53	7	6	4
21	6202120002106	BAAMANG TENGAH	122	19	103	10	34	33	26
22	6202120003015	BAAMANG HULU	138	3	135	34	42	42	17
23	6202190003008	CEMPAKA MULIA BARAT	121	40	81	55	2	10	14
24	6202190003012		241	79	162	122	10	18	12
25	6202211004003	TUMBANG PENYAHUAN	149	13	136	122	0	9	5
26	6203040010019	MURUNG KERAMAT	106	9	97	93	0	3	1
27	6203040011027	SELAT HILIR	122	19	103	2	6	28	67
28	6203040013013	SELAT HULU	181	34	147	28	28	64	27
29	6203040015004	PULAU TELO	137	15	122	37	16	23	46
30	6203041005006	TERUSAN RAYA BARAT	49	4	45	45	0	0	0
31	6203041012014	PULAU KUPANG	226	16	210	173	15	12	10
32	6203041013004	SEI LUNUK	208	46	162	37	76	37	12
33	6203080001002	MAMBULAU	172	72	100	9	26	43	22
34	6203080004001	BARIMBA	184	38	146	65	20	24	37

No	kode_bs	Desa/ Kelurahan	X1	X2	X3	X4	X5	X6	X7
35	6203100003010	PALINGKAU LAMA	176	30	146	45	4	81	16
36	6203110005002	PANTAI	144	38	106	98	0	0	8
37	6203150019001	MANTANGAI HILIR	147	16	131	118	5	6	2
38	6203150019011		177	33	144	88	21	24	11
39	6203150026003	KATUNJUNG	163	28	135	126	6	0	3
40	6203170006004	PUJON	774	212	562	19	139	333	71
41	6203180018001	JAKATAN PARI	266	67	199	140	38	16	5
42	6204020006004	DAMPARAN	256	39	217	189	0	18	10
43	6204030002014	BANGKUANG	116	10	106	59	17	16	14
44	6204030004002	JANGGI	225	82	143	97	22	10	14
45	6204040004001	DANAU SADAR	301	42	259	143	86	14	16
46	6204040006016	HILIR SPER	226	54	172	1	23	61	87
47	6204040014047	BUNTOK KOTA	145	18	127	19	19	14	75
48	6204040026007	TANJUNG JAWA	79	8	71	67	1	0	3
49	6205010002009	TUMPUNG LAUNG II	186	35	151	146	0	0	5
50	6205050012017	LANJAS	209	41	168	3	20	35	110
51	6205050016009	BUKIT SAWIT	197	36	161	71	3	54	33
52	6206011003005	SUNGAI CABANG BARAT	186	51	135	97	16	9	13
53	6206020004003	PADANG	300	52	248	47	43	82	76
54	6207010001003	BATU KOTAM	293	98	195	141	5	35	14
55	6207010005008	NANGA BULIK	268	90	178	29	42	46	61
56	6207020010004	TAPIN BINI	176	20	156	88	8	12	48
57	6207021006007	KARANG BESI	136	26	110	102	1	2	5
58	6208010005002	KUALA PEMBUANG SATU	237	46	191	5	46	46	94
59	6208010005014		168	16	152	40	53	24	35
60	6208020012006	TABIKU	95	19	76	70	2	2	2
61	6208030007014	ASAM BARU	226	24	202	82	37	51	32
62	6208040010011	DURIAN KAIT	169	38	131	126	0	3	2
63	6208040026001	RANTAU PULUT	95	25	70	36	21	7	6
64	6209030003003	TELAGA	375	125	250	99	70	64	17
65	6209050004013	KASONGAN LAMA	221	51	170	20	25	43	82
66	6209050004022		379	87	292	27	40	50	175
67	6209050008035	HAMPALIT	111	43	68	3	31	22	12
68	6209060004003	BANGKUANG	183	10	173	83	64	15	11
69	6209060006003	PENDAHARA	155	21	134	77	19	2	36
70	6209060010002	TEWANG MANYANGEN	106	10	96	90	2	2	2
71	6209080001002	MIRAH KALANAMAN	121	48	73	57	1	2	13

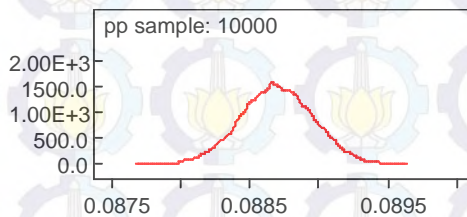
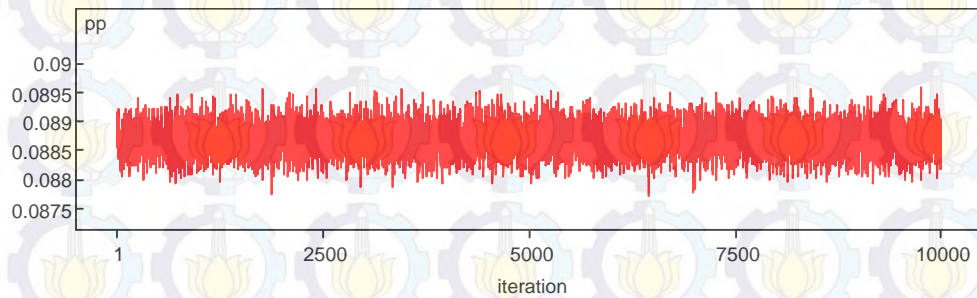
No	kode_bs	Desa/ Kelurahan	X1	X2	X3	X4	X5	X6	X7
72	6209090004011	TUMBANG MANGGO	111	21	90	60	10	7	13
73	6209090011001	TUMBANG KAWEI	111	39	72	22	41	5	4
74	6209100006002	TUMBANG HIRAN	134	17	117	69	38	5	5
75	6209100010003	TUMBANG BEMBAN	81	41	40	13	24	0	3
76	6209100013002	TUMBANG TAEI	85	13	72	42	22	1	7
77	6209111009003	TUMBANG KARUE	122	62	60	60	0	0	0
78	6210010003002	KIAPAK	210	43	167	147	2	10	8
79	6210010008008	BAHAUR TENGAH	201	18	183	142	12	13	16
80	6210011002002	PADURAN MULIA	95	4	91	85	1	5	0
81	6210020015008	KANTAN ATAS	200	9	191	171	2	11	7
82	6210040003008	MANTAREN II	240	50	190	47	84	42	17
83	6210040007003	GOHONG	104	10	94	66	11	1	16
84	6210040010006	BERENG	102	11	91	66	3	14	8
85	6210040010007		162	8	154	106	9	22	17
86	6210050003002	BUKIT RAWI	229	22	207	120	12	9	66
87	6210060011001	TUMBANG TARUSAN	151	7	144	85	5	10	44
88	6211010001004	TAKARAS	184	96	88	44	13	11	20
89	6211030003008	SEPANG KOTA	155	49	106	36	40	15	15
90	6211040012018	KUALA KURUN	218	12	206	92	76	20	18
91	6211060008002	TUMBANG MIRI	271	40	231	173	32	9	17
92	6211060011002	TUMBANG TAJUNGAN	255	8	247	147	82	5	13
93	6212050007002	PUTAI	177	2	175	141	23	3	8
94	6212050009003	NETAMPIN	201	10	191	93	17	46	35
95	6213010011001	PANTAI LAGA	224	72	152	68	62	8	14
96	6213020011023	BERIWIT	75	10	65	27	6	4	28
97	6213030003009	MUARA LAUNG I	152	12	140	75	33	23	9
98	6213040032001	KARALI	360	15	345	252	32	18	43
99	6213052003002	KALASIN	310	72	238	207	11	2	18
100	6271010004062	PAHANDUT	315	72	243	17	54	133	39
101	6271010004068		164	30	134	1	2	84	47
102	6271010006012	LANGKAI	160	17	143	2	26	28	87
103	6271010006053		305	91	214	3	30	99	82
104	6271010011010	PAHANDUT SEBERANG	95	35	60	15	13	24	8
105	6271012001009	MENTENG	232	22	210	5	72	71	62
106	6271012001028		290	65	225	5	54	46	120
107	6271012001064		160	33	127	0	24	31	72

No	kode_bs	Desa/ Kelurahan	X1	X2	X3	X4	X5	X6	X7
108	6271012002004	PALANGKA	118	35	83	2	3	17	61
109	6271012002038		56	22	34	2	4	10	18
110	6271012002126		180	56	124	8	58	33	25
111	6271012003021	BUKIT TUNGAL	156	12	144	0	29	47	68
112	6271012003060		232	27	205	1	54	69	81
113	6271012003069		130	27	103	11	25	29	38
114	6271012003075		184	11	173	15	23	25	110
115	6271012004001	PETUK KATIMPUN	73	14	59	44	3	7	5
116	6271021001005	PETUK BUKIT	74	20	54	20	16	4	14

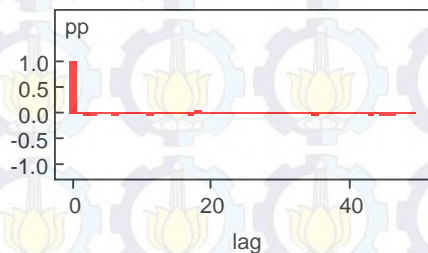
Lampiran 9. Syntax Program WinBUGS dan Hasil dari Estimasi Parameter Proporsi Pengangguran dari Sampel BPS

```
# Model
model; {
  for (i in 1 : s) {
    x[i] ~ dbin(pp, n[i]) # n=X1, x=X2
  }
  pp ~ dbeta(101350,1044483) # Prior sensus penduduk
}

# Data
list(x= c(0, 19, 1, 14, 26, 108, 1, 13, 0, 24, 35, 54, 0, 108, 22, 47, 3, 8, 25, 3, 29, 0, 0, 7, 18, 16, 10, 0, 4,
0, 1, 9, 0, 14, 30, 0, 21, 11, 7, 20, 0, 13, 41, 39, 52, 6, 0, 2, 0, 1, 50, 6, 9, 3, 16, 12, 64, 59, 10, 19, 65,
8, 33, 50, 0, 24, 8, 43, 50, 2, 25, 2, 29, 17, 12, 34, 27, 18, 8, 6, 0, 3, 8, 4, 4, 0, 4, 0, 1, 6, 17, 14, 7, 18,
9, 0, 1, 31, 4, 41, 33, 0, 9, 4, 8, 0, 54, 105, 30, 35, 10, 40, 3, 8, 8, 8),
n= c(237, 114, 115, 201, 295, 289, 60, 242, 171, 124, 126, 229, 191, 400, 215, 132, 190, 177, 194, 92, 234,
137, 172, 158, 125, 120, 98, 204, 169, 142, 153, 117, 83, 106, 133, 103, 249, 143, 139, 122, 115, 202, 222,
170, 264, 71, 171, 172, 111, 113, 232, 104, 121, 210, 227, 170, 190, 207, 138, 249, 219, 152, 139, 226, 118,
265, 92, 206, 94, 152, 148, 114, 155, 148, 108, 202, 136, 98, 110, 137, 104, 194, 170, 286, 113, 220, 228, 164,
174, 107, 289, 121, 155, 129, 318, 188, 149, 276, 98, 121, 293, 189, 167, 173, 668, 261, 154, 181, 295, 246,
73, 227, 131, 186, 191, 141), s=116) # observed data
# Inits
list(pp = 0.05) # starting values for pp
```

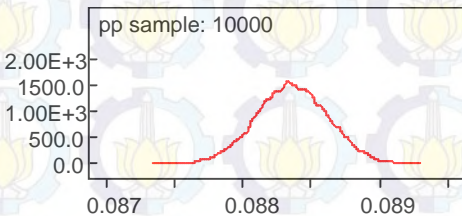
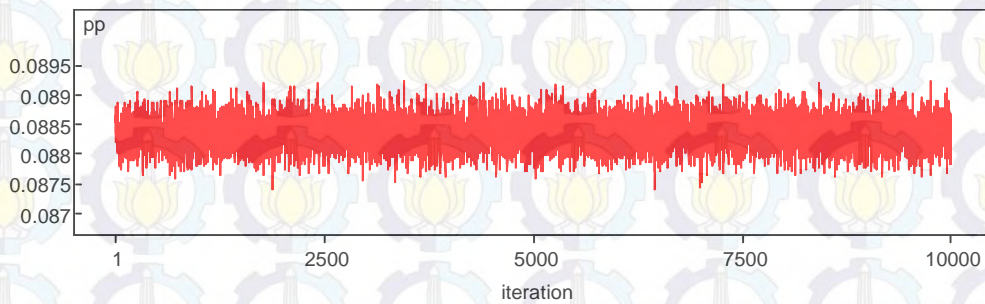


node	mean	sd	MC error	2.5%	median	97.5%	start	sample
pp	0.0887	2.619E-4	2.873E-6	0.08819	0.0887	0.08922	1	10000

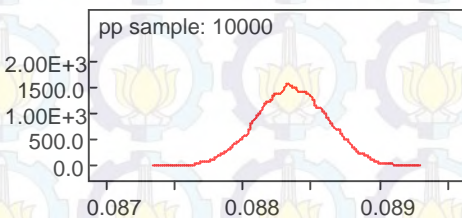


Lampiran 10. Syntax Program WinBUGS dan Hasil dari Estimasi Parameter Proporsi Pengangguran dari Sampel Hasil *Cube Method*

```
# Model
model; {
for (i in 1 : s) {
x[i] ~ dbin(pp, n[i]) # n=X1, x=X2
}
pp ~ dbeta(101350,1044483) # Prior sensus penduduk
}
# Data
list(x= c(18, 21, 20, 42, 21, 27, 8, 4, 70, 1, 2, 0, 7, 0, 11, 4, 7, 8, 3, 66, 9, 34, 33, 2, 0, 5, 10, 3, 5, 2, 0, 3, 1, 2, 0,
7, 11, 7, 0, 12, 0, 20, 15, 18, 1, 3, 17, 6, 0, 0, 0, 8, 40, 9, 18, 10, 16, 12, 22, 46, 7, 0, 17, 3, 22, 1, 10, 3, 8, 46, 1,
0, 4, 7, 49, 20, 5, 11, 30, 30, 15, 17, 12, 36, 1, 54, 0, 1, 6, 0, 6, 9, 38, 15, 1, 26, 65, 22, 2, 0, 7, 4, 0, 44, 15, 12,
32, 22, 26, 4, 57, 3, 14, 20, 66, 28),
n= c(221, 225, 194, 252, 284, 171, 136, 167, 231, 139, 148, 273, 203, 123, 202, 142, 143, 134, 310, 290, 107,
129, 114, 201, 124, 144, 163, 138, 160, 166, 133, 171, 188, 162, 115, 132, 450, 85, 353, 206, 61, 192, 176,
165, 104, 151, 188, 109, 197, 290, 91, 134, 154, 77, 151, 116, 84, 164, 204, 143, 88, 183, 86, 377, 138, 102,
134, 143, 91, 237, 138, 86, 171, 236, 128, 195, 200, 250, 140, 237, 152, 100, 47, 250, 207, 184, 201, 136, 186,
286, 91, 192, 145, 124, 232, 274, 209, 185, 182, 69, 150, 180, 163, 238, 360, 159, 188, 165, 89, 290, 246, 54,
257, 181, 257, 149), s=116) # observed data
# Inits
list(pp = 0.05) # starting values for pp
```



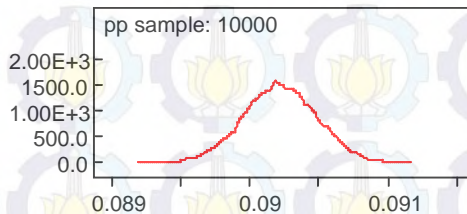
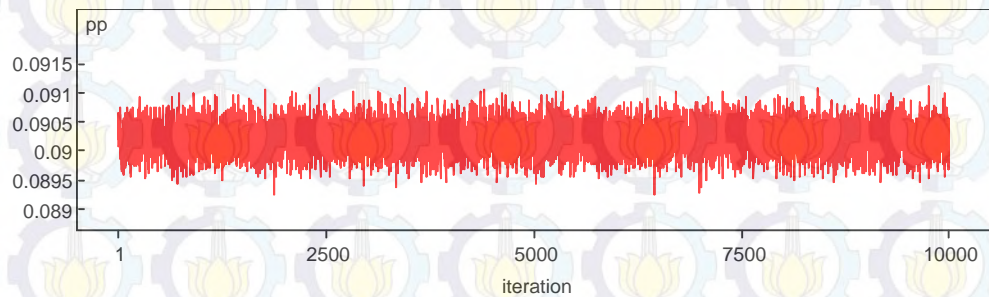
node	mean	sd	MC error2.5%	median	97.5%	start	sample
pp	0.08836	2.614E-4	2.869E-6	0.08836	0.08888	1	10000



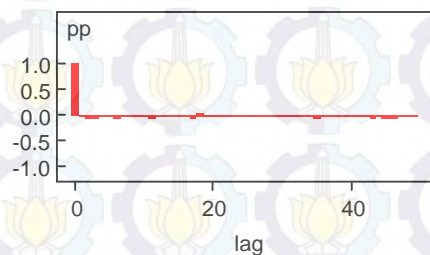
Lampiran 11. Syntax Program WinBUGS dan Hasil dari Estimasi Parameter Proporsi Pengangguran dari Sampel Hasil *Reference Prior*

```
# Model
model; {
  for (i in 1 : s) {
    x[i] ~ dbin(pp, n[i]) # n=X1, x=X2
  }
  pp ~ dbeta(101350,1044483) # Prior sensus penduduk
}

# Data
list(x= c(57, 34, 22, 108, 62, 25, 49, 15, 22, 32, 2, 91, 2, 32, 105, 4, 8, 108, 66, 36, 19, 3, 40, 79, 13, 9, 19, 34,
15, 4, 16, 46, 72, 38, 30, 38, 16, 33, 28, 212, 67, 39, 10, 82, 42, 54, 18, 8, 35, 41, 36, 51, 52, 98, 90,
20, 26, 46, 16, 19, 24, 38, 25, 125, 51, 87, 43, 10, 21, 10, 48, 21, 39, 17, 41, 13, 62, 43, 18, 4, 9,
50, 10, 11, 8, 22, 7, 96, 49, 12, 40, 8, 2, 10, 72, 10, 12, 15, 72, 72, 30, 17, 91, 35, 22, 65, 33, 35, 22, 56, 12, 27,
27, 11, 14, 20),
n= c(253, 130, 205, 289, 554, 175, 191, 173, 205, 229, 252, 177, 175, 138, 213, 280, 172, 400, 290, 106, 122,
138, 121, 241, 149, 106, 122, 181, 137, 49, 226, 208, 172, 184, 176, 144, 147, 177, 163, 774, 266, 256,
116, 225, 301, 226, 145, 79, 186, 209, 197, 186, 300, 293, 268, 176, 136, 237, 168, 95, 226, 169, 95,
375, 221, 379, 111, 183, 155, 106, 121, 111, 111, 134, 81, 85, 122, 210, 201, 95, 200, 240, 104, 102, 162, 229,
151, 184, 155, 218, 271, 255, 177, 201, 224, 75, 152, 360, 310, 315, 164, 160, 305, 95, 232, 290, 160,
118, 56, 180, 156, 232, 130, 184, 73, 74), s=116) # observed data
# Inits
list(pp = 0.05) # starting values for pp
```



node	mean	sd	MC error	2.5%	median	97.5%	start	sample
pp	0.09022	2.639E-4	2.897E-6	0.0897	0.09022	0.09075	1	10000



Lampiran 12. Syntax Program WinBUGS Estimasi Parameter Proporsi Pekerja menurut Lapangan Usaha dari Sampel Sakernas BPS

```

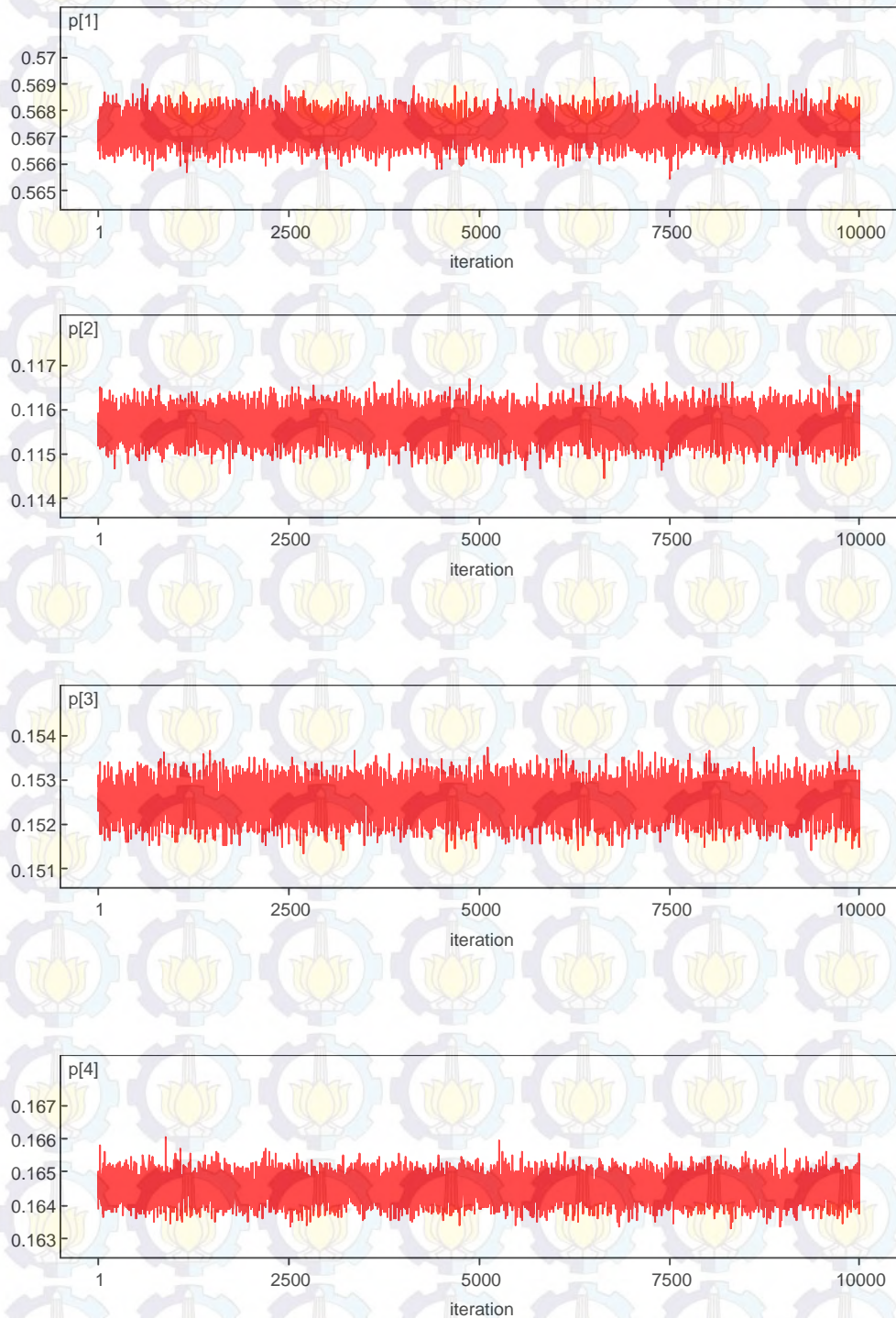
model;
{
for (i in 1:S) {
x[i,1:4] ~ dmulti(p[1:4], n[i])# Likelihood
}
p[1:4] ~ ddirch(alpha[]) #Prior
}

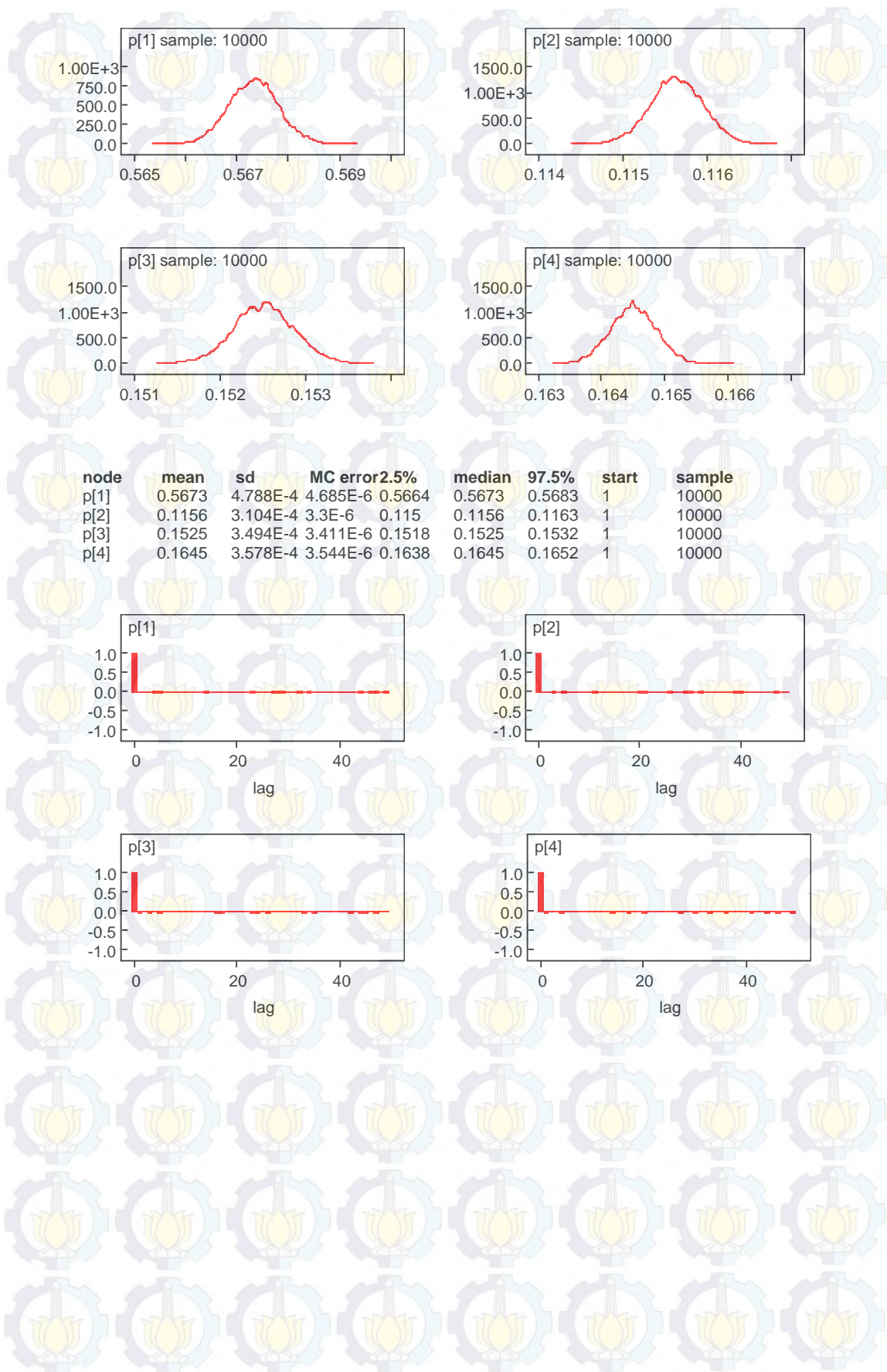
# Data
list(x= structure(.Data= c(167, 37, 23, 16, 17, 28, 54, 29, 167, 33, 72, 130, 39, 28, 15, 81, 168, 14, 24, 33, 172,
121, 168, 145, 85, 23, 73, 198, 121, 67, 136, 77, 63, 67, 86, 78, 8, 1, 1, 90, 96, 140, 127, 10, 6, 4, 166, 133,
111, 67, 28, 15, 74, 156, 13, 98, 29, 112, 50, 206, 115, 9, 15, 155, 110, 123, 61, 117, 43, 143, 66, 10, 14, 2, 51,
100, 56, 75, 82, 79, 72, 18, 32, 213, 105, 128, 224, 156, 148, 61, 193, 11, 126, 10, 233, 164, 120, 194, 56, 63,
174, 177, 106, 0, 22, 175, 43, 0, 163, 11, 0, 10, 4, 4, 31, 66, 17, 27, 35, 14, 40, 106, 0, 33, 2, 43, 7, 11, 51, 64,
32, 0, 4, 32, 12, 23, 6, 2, 0, 0, 7, 7, 5, 1, 3, 6, 4, 5, 7, 0, 6, 5, 32, 14, 40, 3, 5, 4, 37, 10, 16, 19, 0, 12, 0, 17, 50,
21, 18, 2, 14, 4, 21, 7, 4, 3, 1, 14, 24, 1, 0, 45, 9, 14, 1, 0, 12, 10, 37, 29, 39, 41, 27, 1, 6, 28, 19, 30, 26, 39, 2,
24, 0, 0, 2, 19, 18, 28, 11, 30, 17, 2, 2, 15, 18, 12, 39, 0, 1, 13, 88, 4, 35, 13, 38, 32, 14, 23, 43, 25, 59, 45,
39, 27, 39, 97, 103, 27, 4, 71, 2, 6, 8, 17, 48, 116, 41, 3, 5, 56, 89, 17, 17, 4, 4, 3, 9, 29, 9, 4, 16, 15, 9, 5, 5, 10,
6, 12, 36, 97, 18, 4, 9, 14, 7, 31, 110, 17, 5, 19, 0, 17, 42, 34, 12, 36, 45, 39, 35, 19, 45, 9, 17, 31, 18, 12, 6, 49,
9, 17, 0, 2, 6, 15, 40, 85, 3, 19, 9, 2, 4, 8, 9, 45, 22, 12, 1, 14, 0, 5, 12, 5, 38, 34, 2, 29, 42, 13, 17, 18, 6, 3, 40,
12, 27, 86, 100, 65, 8, 45, 33, 35, 19, 81, 54, 28, 72, 12, 14, 4, 17, 60, 109, 20, 1, 96, 0, 18, 4, 17, 53, 84, 105,
1, 10, 67, 44, 16, 10, 10, 0, 3, 6, 45, 1, 1, 25, 54, 3, 21, 8, 15, 5, 8, 152, 20, 73, 5, 5, 31, 10, 80, 80, 25, 0, 6, 0,
11, 62, 28, 8, 13, 139, 17, 41, 10, 29, 12, 21, 90, 49, 8, 2, 24, 5, 15, 0, 5, 39, 77, 35, 15, 3, 8, 17, 2, 10, 16, 4,
98, 82, 18, 1, 54, 0, 3, 11, 16, 23, 34, 9, 42, 17, 9, 9, 18, 14, 2, 7, 0, 24, 70, 450, 17, 14, 18, 31, 133, 30, 73, 27,
121, 21, 10), .Dim= c(116,4)),
n= c(237, 95, 114, 187, 269, 181, 59, 229, 171, 100, 91, 175, 191, 292, 193, 85, 187, 169, 169, 89, 205, 137,
172, 151, 107, 104, 88, 204, 165, 142, 152, 108, 83, 92, 103, 103, 228, 132, 132, 102, 115, 189, 181, 131, 212,
65, 171, 170, 111, 112, 182, 98, 112, 207, 211, 158, 126, 148, 128, 230, 154, 144, 106, 176, 118, 241, 84, 163,
44, 150, 123, 112, 126, 131, 96, 168, 109, 80, 102, 131, 104, 191, 162, 282, 109, 220, 224, 164, 173, 101, 272,
107, 148, 111, 309, 188, 148, 245, 94, 80, 260, 189, 158, 169, 660, 261, 100, 76, 265, 211, 63, 187, 128, 178,
183, 133), S=116, alpha=c(597862,118288,157868,170465))

#Initial values
list(p=c(0.25,0.25,0.25,0.25))

```


Lampiran 13. Hasil Estimasi Parameter Proporsi Pekerja menurut Lapangan Usaha dari Sampel Sakernas BPS





Lampiran 14. Syntax Program WinBUGS Estimasi Parameter Proporsi Pekerja menurut Lapangan Usaha dari Sampel Hasil *Cube Method*

```

model;
{
  for (i in 1:S) {

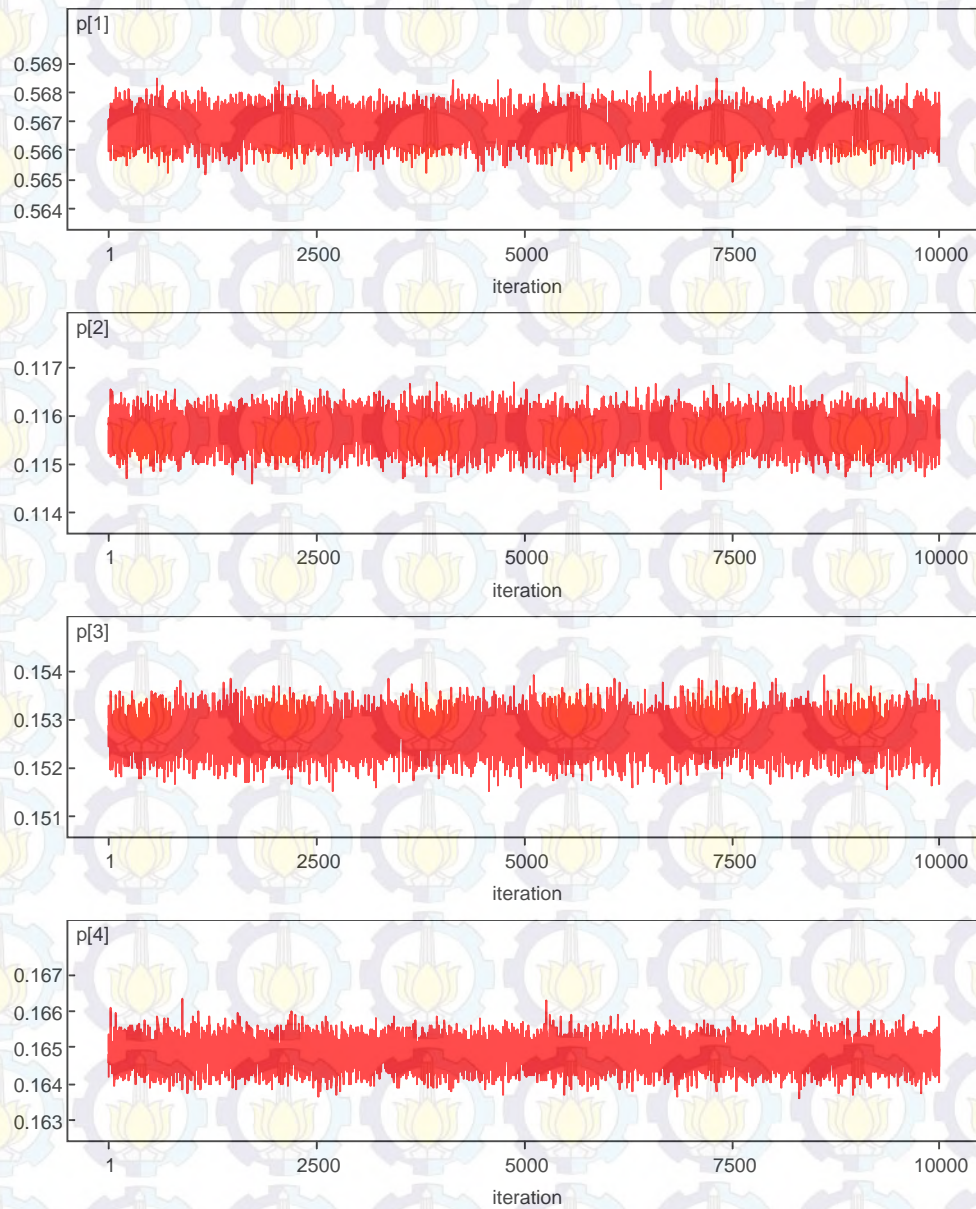
    x[i,1:4] ~ dmulti(p[1:4], n[i])# Likelihood x=X4,X5,X6,X7 n=x3
  }
  p[1:4] ~ ddirch(alpha[]) #Prior
}

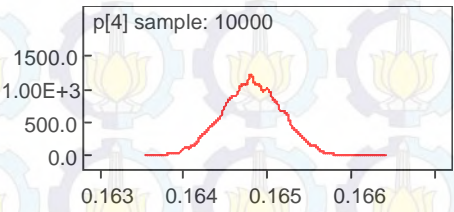
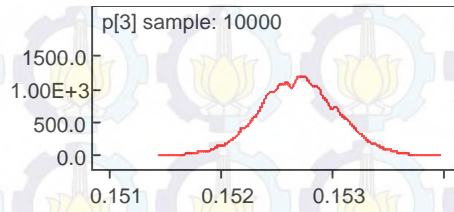
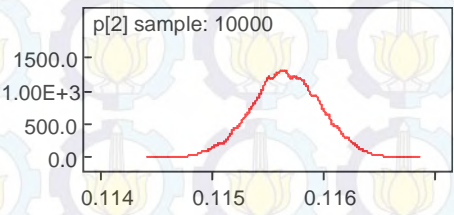
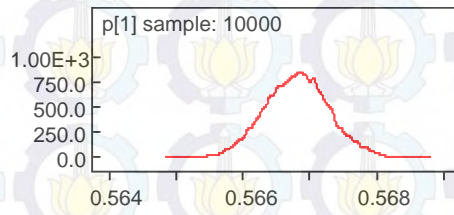
# Data
list(x= structure(.Data= c(165, 140, 113, 4, 20, 16, 40, 3, 13, 33, 143, 244, 146, 63, 113, 130, 13, 7, 12, 46, 78,
69, 37, 115, 75, 14, 22, 34, 145, 146, 122, 107, 163, 120, 84, 51, 376, 49, 342, 171, 60, 109, 14, 4, 39, 109,
142, 57, 192, 240, 91, 83, 26, 55, 105, 59, 31, 73, 120, 70, 57, 161, 0, 358, 28, 85, 51, 106, 68, 5, 78, 71,105,
94, 75, 111, 195, 223, 73, 161, 132, 52, 11, 18, 3, 0, 154, 133, 76, 237, 43, 150, 52, 86, 194, 47, 46, 99, 46, 57,
103, 119, 137, 175, 252, 92, 90, 1, 0, 16, 6, 1, 5, 4, 37, 6, 9, 16, 10, 10, 17, 11, 22, 12, 79, 27, 0, 7, 4, 1, 44, 0,
15, 36, 65, 73, 2, 3, 3, 30, 16, 22, 16, 42, 0, 2, 1, 19, 3, 9, 3, 7, 63, 18, 1, 11, 0, 10, 21, 9, 20, 5, 1, 12, 1, 42, 0,
1, 75, 0, 6, 17, 26, 54, 25, 12, 6, 0, 5, 0, 35, 1, 4, 1, 1, 46, 3, 7, 8, 17, 0, 30, 0, 7, 26, 15, 0, 24, 3, 77, 11, 74, 6,
0, 15, 11, 2, 17, 5, 15, 21, 81, 64, 24, 118, 1, 3, 21, 1, 4, 32, 21, 39, 10, 10, 48, 34, 11, 25, 43, 6, 56, 8, 38, 31,
61, 142, 70, 32, 124, 42, 50, 0, 11, 34, 23, 19, 1, 71, 49, 138, 36, 11, 8, 11, 28, 19, 52, 77, 42, 6, 9, 7, 10, 11,
17, 18, 50, 0, 5, 1, 7, 1, 25, 93, 23, 25, 27, 16, 21, 1, 5, 0, 14, 12, 6, 13, 16, 6, 17, 26, 5, 6, 17, 25, 2, 31, 1, 19,
10, 2, 46, 7, 4, 18, 20, 3, 16, 0, 9, 7, 13, 1, 3, 7, 78, 170, 47, 12, 0, 63, 17, 24, 10, 23, 4, 10, 33, 20, 8, 11, 3, 8,
17, 14, 3, 18, 7, 9, 51, 33, 51, 63, 9, 40, 75, 32, 31, 21, 10, 20, 135, 84, 47, 34, 24, 27, 28, 3, 11, 12, 36, 15, 7,
37, 34, 92, 69, 7, 15, 30, 26, 14, 51, 38, 17, 4, 7, 3, 32, 10, 14, 10, 17, 0, 6, 9, 5, 0, 28, 33, 111, 19, 7, 12, 13, 3,
3, 0, 28, 1, 7, 9, 14, 5, 8, 11, 10, 12, 5, 39, 14, 22, 14, 50, 23, 12, 94, 49, 4, 36, 98, 1, 18, 0, 0, 4, 18, 4, 4, 14,
41, 22, 9, 29, 2, 26, 21, 16, 6, 27, 4, 6, 87, 14, 32, 5, 8, 29, 19, 11, 12, 43, 27, 18, 81, 20, 171, 86, 30, 173, 39,
116, 28), .Dim= c(116,4)),
n= c(203, 204, 174, 210, 263, 144, 128, 163, 161, 138, 146, 273, 196, 123, 191, 138, 136, 126, 307, 224, 98,
95, 81, 199, 124, 139, 153, 135, 155, 164, 133, 168, 187, 160, 115, 125, 439, 78, 353, 194, 61, 172, 161, 147,
103, 148, 171, 103, 197, 290, 91, 126, 114, 68, 133, 106, 68, 152, 182, 97, 81, 183, 69, 374, 116, 101, 124,
140, 83, 191, 137, 86, 167, 229, 79, 175, 195, 239, 110, 207, 137, 83, 35, 214, 206, 130, 201, 135, 180, 286,
85, 183, 107, 109, 231, 248, 144, 63, 180, 69, 143, 176, 163, 194, 345, 147, 156, 143, 63, 286, 189, 51, 243,
161, 191, 121), S=116, alpha=c(597862,118288,157868,170465))

#Initial values
list(p=c(0.25,0.25,0.25,0.25))

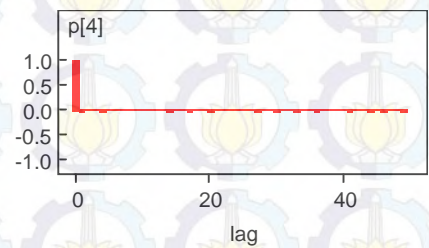
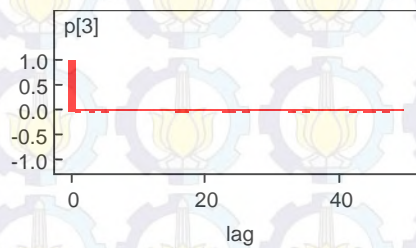
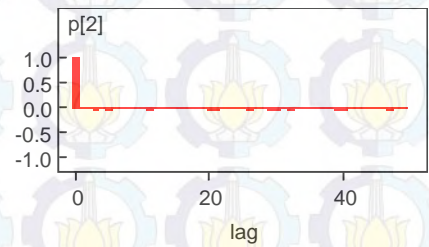
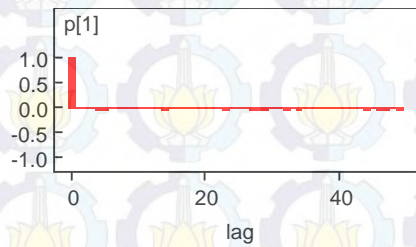
```


Lampiran 15. Hasil Estimasi Parameter Proporsi Pekerja menurut Lapangan Usaha dari Sampel Hasil *Cube Method*





node	mean	sd	MC error	2.5%	median	97.5%	start	sample
p[1]	0.5668	4.788E-4	4.685E-6	0.5659	0.5668	0.5678	1	10000
p[2]	0.1157	3.103E-4	3.3E-6	0.1151	0.1157	0.1163	1	10000
p[3]	0.1527	3.495E-4	3.412E-6	0.152	0.1527	0.1534	1	10000
p[4]	0.1648	3.58E-4	3.547E-6	0.1641	0.1648	0.1655	1	10000



Lampiran 16. Syntax Program WinBUGS Estimasi Parameter Proporsi Pekerja menurut Lapangan Usaha dari Sampel Hasil *Reference Prior*

```

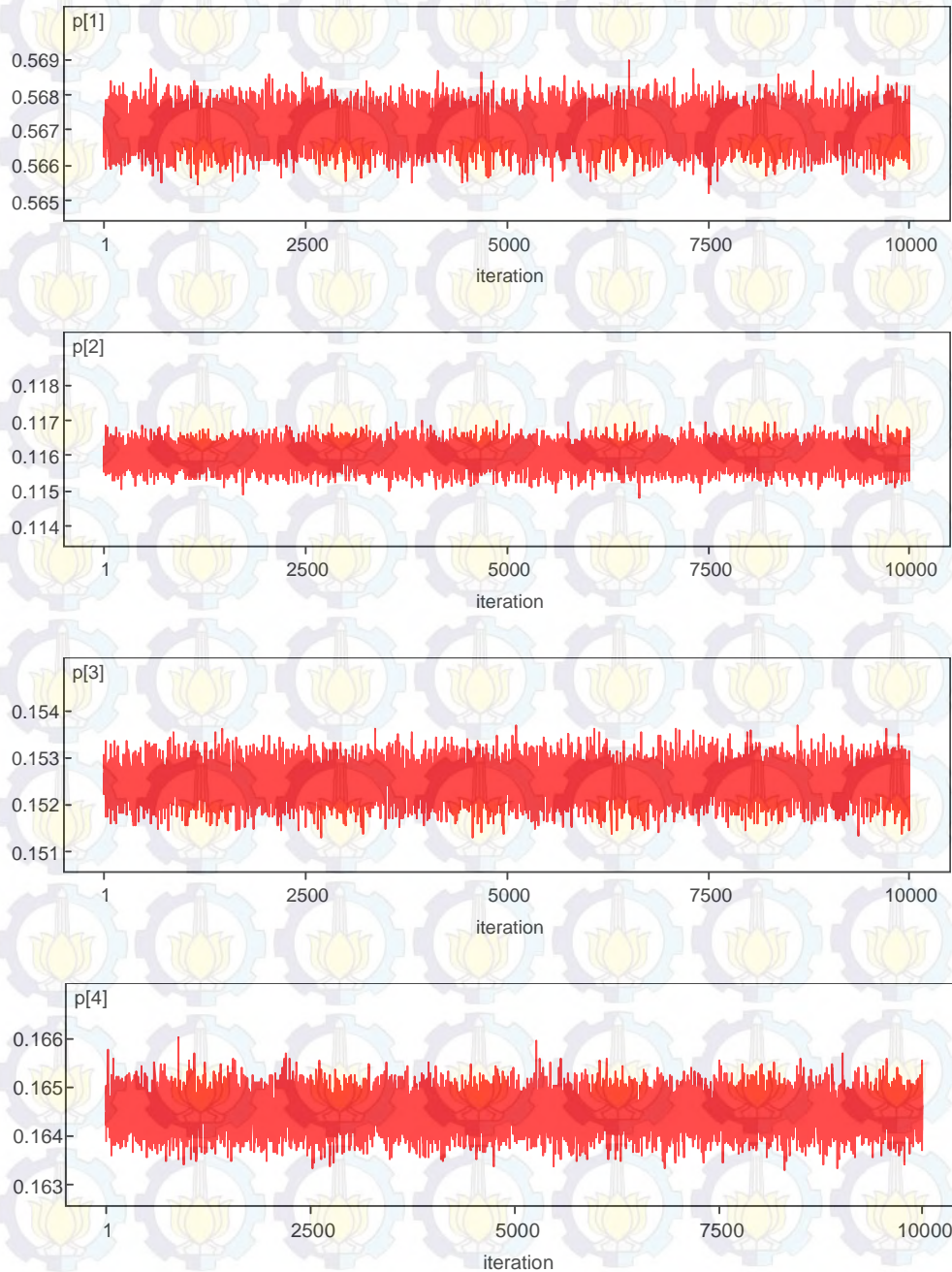
model;
{
for (i in 1:S) {
x[i,1:4] ~ dmulti(p[1:4], n[i])# Likelihood
}
p[1:4] ~ ddirch(alpha[]) #Prior informative dari sensus penduduk
}

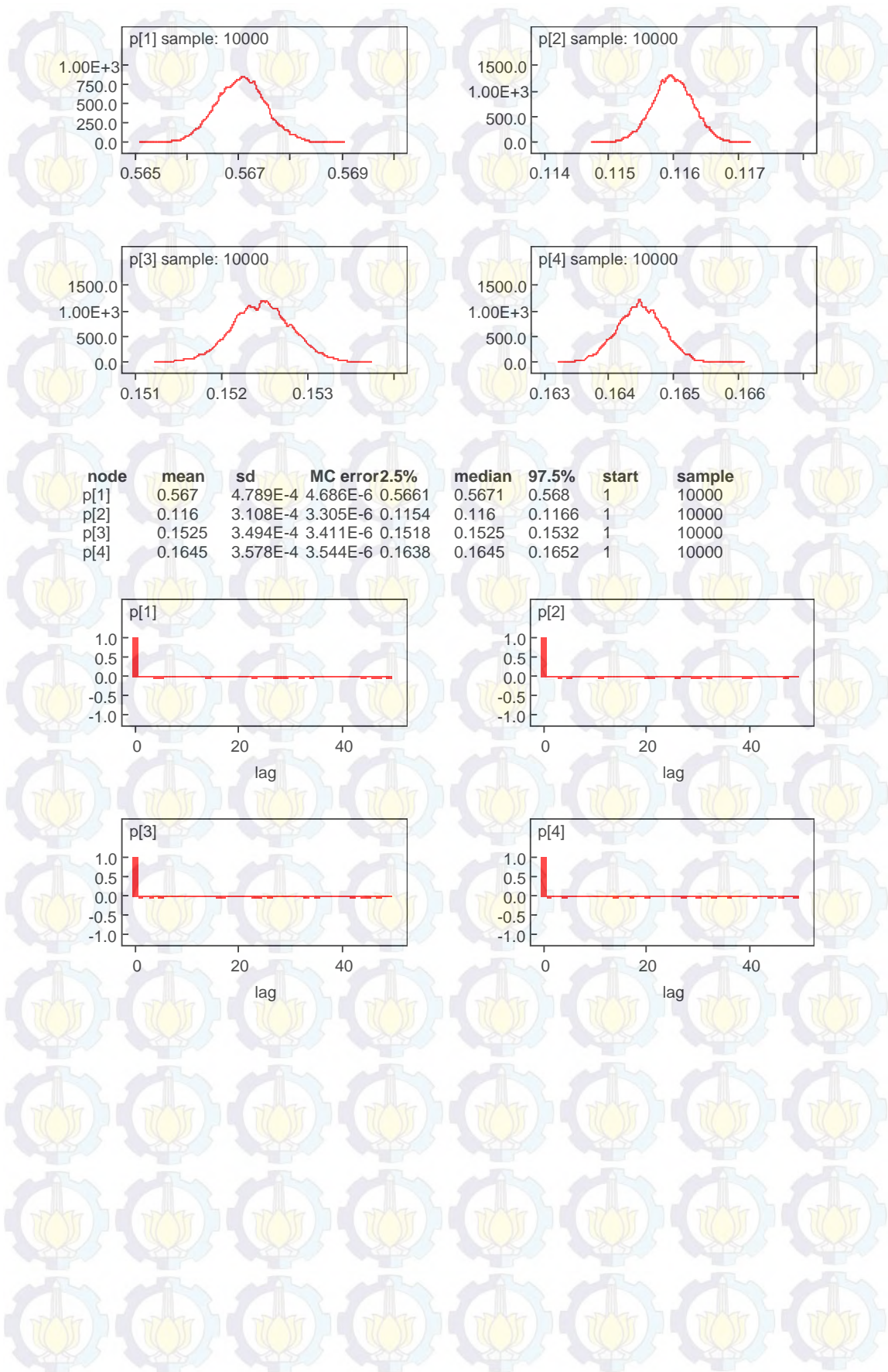
# Data
list(x= structure(.Data= c(77, 4, 147, 28, 492, 96, 28, 25, 54, 175, 248, 41, 160, 85, 71, 43, 19, 28, 46, 53, 10,
34, 55, 122, 122, 93, 2, 28, 37, 45, 173, 37, 9, 65, 45, 98, 118, 88, 126, 19, 140, 189, 59, 97, 143, 1, 19,
67, 146, 3, 71, 97, 47, 141, 29, 88, 102, 5, 40, 70, 82, 126, 36, 99, 20, 27, 3, 83, 77, 90, 57, 60, 22, 69, 13,
42, 60, 147, 142, 85, 171, 47, 66, 66, 106, 120, 85, 44, 36, 92, 173, 147, 141, 93, 68, 27, 75, 252, 207,
17, 1, 2, 3, 15, 5, 5, 0, 2, 2, 8, 0, 1, 11, 15, 44, 20, 17, 5, 12, 106, 0, 40, 45, 80, 40, 1, 0, 22, 0, 8, 21, 58, 33,
64, 73, 7, 34, 42, 2, 10, 0, 0, 6, 28, 16, 0, 15, 76, 26, 20, 4, 0, 5, 21, 6, 139, 38, 0, 17, 22, 86, 23, 19, 1, 0, 20,
3, 16, 43, 5, 42, 8, 1, 46, 53, 2, 37, 0, 21, 70, 25, 40, 31, 64, 19, 2, 1, 10, 41, 38, 24, 22, 0, 2, 12, 1, 2, 84, 11, 3,
9, 12, 5, 13, 40, 76, 32, 82, 23, 17, 62, 6, 33, 32, 11, 54, 2, 26, 30, 13, 72, 54, 24, 3, 4, 58, 29, 54, 25, 23, 3,
16, 64, 68, 15, 27, 0, 13, 37, 38, 68, 2, 2, 18, 0, 4, 7, 141, 60, 116, 36, 6, 33, 42, 10, 18, 9, 3, 28, 64, 23, 0,
12, 37, 43, 24, 81, 0, 6, 24, 0, 333, 16, 18, 16, 10, 14, 61, 14, 0, 0, 35, 54, 9, 82, 35, 46, 12, 2, 46, 24, 2, 51,
3, 7, 64, 43, 50, 22, 15, 2, 2, 2, 7, 5, 5, 0, 1, 0, 10, 13, 5, 11, 42, 1, 14, 22, 9, 10, 11, 15, 20, 9, 5, 3, 46, 8, 4, 23,
18, 2, 133, 84, 28, 99, 24, 71, 46, 31, 17, 10, 33, 47, 69, 29, 25, 7, 4, 38, 19, 9, 20, 0, 1, 32, 15, 21, 19, 0, 5, 13,
9, 9, 34, 52, 84, 69, 4, 26, 17, 14, 12, 5, 1, 67, 27, 46, 0, 10, 12, 22, 37, 16, 8, 2, 11, 3, 71, 5, 10, 14, 14,
16, 87, 75, 3, 5, 110, 33, 13, 76, 14, 61, 48, 5, 94, 35, 2, 32, 2, 6, 17, 82, 175, 12, 11, 36, 2, 13, 13, 4, 5, 3, 7, 0,
8, 16, 0, 7, 17, 16, 8, 17, 66, 44, 20, 15, 18, 17, 13, 8, 35, 14, 28, 9, 43, 18, 39, 47, 87, 82, 8, 62, 120, 72, 61,
18, 25, 68, 81, 38, 110, 5, 14), .Dim= c(116,4)),
n= c(196, 96, 183, 181, 492, 150, 142, 158, 183, 197, 250, 86, 173, 106, 108, 276, 164, 292, 224, 70, 103, 135,
81, 162, 136, 97, 103, 147, 122, 45, 210, 162, 100, 146, 146, 106, 131, 144, 135, 562, 199, 217, 106, 143, 259,
172, 127, 71, 151, 168, 161, 135, 248, 195, 178, 156, 110, 191, 152, 76, 202, 131, 70, 250, 170, 292, 68, 173,
134, 96, 73, 90, 72, 117, 40, 72, 60, 167, 183, 91, 191, , 190, 94, 91, 154, 207, 144, 88, 106, 206, 231, 247,
175, 191, 152, 65, 140, 345, 238, 243, 134, 143, 214, 60, 210, 225, 127, 83, 34, 124, 144, 205, 103, 173, 59,
54), S=116, alpha=c(597862,118288,157868,170465))

#Initial values
list(p=c(0.25,0.25,0.25,0.25))

```


Lampiran 17. Hasil Estimasi Parameter Proporsi Pekerja menurut Lapangan Usaha dari Sampel Hasil *Reference Prior*





BIOGRAFI PENULIS



Penulis dilahirkan di Pekalongan, tepatnya di Desa Waru Kidul, Kecamatan Wiradesa, Kabupaten Pekalongan, Jawa Tengah pada tanggal 13 Desember 1985. Penulis merupakan putra pertama dari empat bersaudara buah cinta dari pasangan Almarhum Bapak Slamet Sodikin dan Ibu Ulwiyah. Saat ini penulis sudah berkeluarga dengan istri bernama Dyan Ayu Dwi Agustin dengan dua anak, Muhammad Alif Jundullah dan Raisa Saliha Malahati.

Riwayat pendidikan penulis adalah MIS Salafiyah Waru Lor (1992-1998), SLTP Negeri 1 Wiradesa (1998-2001), SMU Negeri 1 Pekalongan (2001-2004), Sekolah Tinggi Ilmu Statistik (STIS) Jakarta (2004-2008). Setelah menamatkan pendidikan DIV di STIS, penulis ditugaskan bekerja di BPS Kabupaten Sukamara Provinsi Kalimantan Tengah sebagai staf Seksi Statistik Sosial. Pada tahun 2013 penulis memperoleh kesempatan beasiswa dari BPS untuk melanjutkan studi S2 di Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA) Institut Teknologi Sepuluh Nopember (ITS) Surabaya.

Surabaya, Maret 2015

Moch. Choeril Anwar
choeril@bps.go.id